

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2004, 16:25:55 ; Search time 44 Seconds  
(without alignments)  
1163.719 Million cell updates/sec

Title: US-10-054-988-114

Perfect score: 962  
Sequence: 1 MEPQLGPEAAALRPGWALL.....DIVQDCHQSQRELKFLCMLR 182

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 28138677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

```
1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	962	100.0	182	9	US-09-739-254-114
2	362	100.0	182	9	US-09-904-615-114
3	962	100.0	182	12	US-10-211-462-217
4	962	100.0	182	14	US-10-054-988-114
5	962	100.0	182	14	US-10-177-293-150
6	962	100.0	182	14	US-10-055-098-114
7	962	100.0	209	9	US-09-739-254-168
8	962	100.0	209	9	US-09-904-615-168
9	962	100.0	209	14	US-10-054-988-168
10	962	100.0	209	14	US-10-055-098-168
11	959	99.7	182	9	US-09-965-528-10
12	959	99.7	182	12	US-10-147-493-172
13	959	99.7	182	12	US-10-145-127-172
14	959	99.7	182	12	US-10-160-503-172
15	959	99.7	182	12	US-10-143-118-172

16	959	99.7	182	12	US-10-144-993-172	Sequence 172, App
17	959	99.7	182	12	US-10-158-787-172	Sequence 172, App
18	959	99.7	182	12	US-09-969-984-10	Sequence 10, Appl
19	959	99.7	182	12	US-10-140-024-172	Sequence 172, App
20	959	99.7	182	12	US-10-140-808-172	Sequence 172, App
21	959	99.7	182	12	US-10-152-405-172	Sequence 172, App
22	959	99.7	182	12	US-10-127-852A-172	Sequence 172, App
23	959	99.7	182	12	US-10-127-900A-172	Sequence 172, App
24	959	99.7	182	12	US-10-128-685A-172	Sequence 172, App
25	959	99.7	182	12	US-10-131-820A-172	Sequence 172, App
26	959	99.7	182	12	US-10-142-886-172	Sequence 172, App
27	959	99.7	182	12	US-10-146-728-172	Sequence 172, App
28	959	99.7	182	12	US-10-146-786-172	Sequence 172, App
29	959	99.7	182	12	US-10-147-499-172	Sequence 172, App
30	959	99.7	182	12	US-10-157-798-172	Sequence 172, App
31	959	99.7	182	14	US-10-121-049-172	Sequence 172, App
32	959	99.7	182	14	US-10-123-904-172	Sequence 172, App
33	959	99.7	182	14	US-10-140-470-172	Sequence 172, App
34	959	99.7	182	14	US-10-175-746-172	Sequence 172, App
35	959	99.7	182	14	US-10-176-918-172	Sequence 172, App
36	959	99.7	182	14	US-10-176-921-172	Sequence 172, App
37	959	99.7	182	14	US-10-137-865-172	Sequence 172, App
38	959	99.7	182	14	US-10-140-474-172	Sequence 172, App
39	959	99.7	182	14	US-10-142-431-172	Sequence 172, App
40	959	99.7	182	14	US-10-143-114-172	Sequence 172, App
41	959	99.7	182	14	US-10-140-002-172	Sequence 172, App
42	959	99.7	182	14	US-10-142-419-172	Sequence 172, App
43	959	99.7	182	14	US-10-123-262-172	Sequence 172, App
44	959	99.7	182	14	US-10-142-423-172	Sequence 172, App
45	959	99.7	182	14	US-10-142-423-172	Sequence 172, App

#### ALIGNMENTS

#### RESULT 1

```
US-09-739-254-114
; Sequence 114, Application US/09739254
; Patent No. US20010021700A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/739,254
; CURRENT FILING DATE: 2000-12-19
; EARLIER APPLICATION NUMBER: 09/511,554
; EARLIER FILING DATE: 2000-02-23
; EARLIER APPLICATION NUMBER: PCT/US99/19330
; EARLIER FILING DATE: 1999-08-24
; EARLIER APPLICATION NUMBER: 60/097,917
; EARLIER FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 60/098,634
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 114
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-739-254-114
```

Query Match 100.0%; Score 962; DB 9; Length 182;  
Best Local Similarity 100.0%; Pred. No. 7.5e-37;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEPQLGPEAAALRPGWALLLWSALSCSFLPASSLSLVPQVRTSNFRTGLDKC	60
Db	1	MEPQLGPEAAALRPGWALLLWSALSCSFLPASSLSLVPQVRTSNFRTGLDKC	60
Qy	61	NACIGTSICKKFKPEEIRSDNWLASHLGLPPDSLLSPANYSDSKIWPRVEIFRLVSKY	120
Db	61	NACIGTSICKKFKPEEIRSDNWLASHLGLPPDSLLSPANYSDSKIWPRVEIFRLVSKY	120

; PRIOR FILING DATE: 2001-02-22

Qy 121 QNEISDRKICASASAPKTCSEIRVLRTKTERFQKWL

; PRIOR FILING DATE: 2001-02-22

Db 121 QNEISDRKICASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOQRELKFLCM 180

QY 181 LR 182  
||

Db 181 LR 182

RESULT 5

US-10-177-293-150

Sequence 150, Application US/10177293

Publication No. US20030124128A1

GENERAL INFORMATION:

APPLICANT: Lillie, James

APPLICANT: Glatt, Karen

APPLICANT: Zhao, Xumei

APPLICANT: Gannavarpu, Manjula

APPLICANT: Kamatkar, Shubhangi

APPLICANT: Mertens, Maureen

APPLICANT: Myer, Vic

APPLICANT: Wang, Youzhen

APPLICANT: Xu, Yongyao

APPLICANT: Hoersch, Sebastian

APPLICANT: Monahan, John

APPLICANT: Meyers, Rachel E.

APPLICANT: Bast Jr., Robert C.

APPLICANT: Hortobagyi, Gabriel N.

APPLICANT: Pusztai, Lajos

APPLICANT: Meric, Funda

APPLICANT: Sahin, Aysegul

APPLICANT: Mills, Gordon B.

TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST CANCER

FILE REFERENCE: MRI-038

CURRENT APPLICATION NUMBER: US/10/177,293

CURRENT FILING DATE: 2002-06-21

PRIOR APPLICATION NUMBER: US 60/299,887

PRIOR FILING DATE: 2001-06-21

PRIOR APPLICATION NUMBER: US 60/301,572

PRIOR FILING DATE: 2001-06-27

PRIOR APPLICATION NUMBER: US 60/306,501

PRIOR FILING DATE: 2001-07-18

PRIOR APPLICATION NUMBER: US 60/325,002

PRIOR FILING DATE: 2001-09-25

PRIOR APPLICATION NUMBER: US 60/362,585

PRIOR FILING DATE: 2002-03-05

PRIOR APPLICATION NUMBER: US 60/xxx,xxx

PRIOR FILING DATE: 2002-05-14

NUMBER OF SEQ ID NOS: 506

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 150

LENGTH: 182

TYPE: PRT

ORGANISM: Homo sapiens

US-10-177-293-150

Query Match 100.0%; Score 962; DB 14; Length 182;

Best Local Similarity 100.0%; Pred. No. 7.5e-97;

Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPLGPEAAALRPGWLALLWVSALSCSFLPASLSLVLPQVRTSYNFGRTFLGLDKC 60

Db 1 MEPLGPEAAALRPGWLALLWVSALSCSFLPASLSLVLPQVRTSYNFGRTFLGLDKC 60

QY 61 NACIGTSICKKFKKEIRSDNWLASHGLPPDSLLSYPNYSDSKIWRRPVEIFRLVSKY 120

Db 61 NACIGTSICKKFKKEIRSDNWLASHGLPPDSLLSYPNYSDSKIWRRPVEIFRLVSKY 120

QY 121 QNEISDRKICASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOQRELKFLCM 180

Db 121 QNEISDRKICASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOQRELKFLCM 180

QY 181 LR 182  
||

Db 181 LR 182

RESULT 6

US-10-055-098-114

Sequence 114, Application US/10055098

Publication No. US20030139954A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 49 Human Secreted Proteins

FILE REFERENCE: P2032P1

CURRENT APPLICATION NUMBER: US/10/055,098

CURRENT FILING DATE: 2002-01-22

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/511,554

PRIOR FILING DATE: EARLIER FILING DATE: 2000-02-23

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US99/19330

PRIOR FILING DATE: EARLIER FILING DATE: 1999-08-24

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/097,917

PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-25

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/098,634

PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-31

NUMBER OF SEQ ID NOS: 170

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 114

LENGTH: 182

TYPE: PRT

ORGANISM: Homo sapiens

US-10-055-098-114

Query Match 100.0%; Score 962; DB 14; Length 182;

Best Local Similarity 100.0%; Pred. No. 7.5e-97;

Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPLGPEAAALRPGWLALLWVSALSCSFLPASLSLVLPQVRTSYNFGRTFLGLDKC 60

Db 1 MEPLGPEAAALRPGWLALLWVSALSCSFLPASLSLVLPQVRTSYNFGRTFLGLDKC 60

QY 61 NACIGTSICKKFKKEIRSDNWLASHGLPPDSLLSYPNYSDSKIWRRPVEIFRLVSKY 120

Db 61 NACIGTSICKKFKKEIRSDNWLASHGLPPDSLLSYPNYSDSKIWRRPVEIFRLVSKY 120

QY 121 QNEISDRKICASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOQRELKFLCM 180

Db 121 QNEISDRKICASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOQRELKFLCM 180

QY 181 LR 182  
||

Db 181 LR 182

RESULT 7

US-09-739-254-168

Sequence 168, Application US/09739254

Patent No. US20010021700A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 49 Human Secreted Proteins

FILE REFERENCE: P2032P1

CURRENT APPLICATION NUMBER: US/09/739,254

CURRENT FILING DATE: 2000-12-19

EARLIER APPLICATION NUMBER: 09/511,554

EARLIER FILING DATE: 2000-02-23

EARLIER APPLICATION NUMBER: PCT/US99/19330

EARLIER FILING DATE: 1999-08-24

EARLIER APPLICATION NUMBER: 60/097,917

EARLIER FILING DATE: 1998-08-25

EARLIER APPLICATION NUMBER: 60/098,634

EARLIER FILING DATE: 1998-08-31

NUMBER OF SEQ ID NOS: 170

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 168

LENGTH: 209

TYPE: PRT

; APPLICANT: Gurney,Austin L.  
; APPLICANT: Sherwood,Steven  
; APPLICANT: Smith,Victoria  
; APPLICANT: Stewart,Timothy A.  
; APPLICANT: Tumas,Daniel  
; APPLICANT: Watanabe,Colin K  
; APPLICANT: Wood,William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C252  
; CURRENT APPLICATION NUMBER: US/10/145,127  
; CURRENT FILING DATE: 2002-05-13  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 172  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-145-127-172

Query Match 99.7%; Score 959; DB 12; Length 182;  
Best Local Similarity 99.5%; Pred. No. 1.6e-96;  
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPOLGPEAAALRPGWLALLLWVSALSCSFSLSLPASSLSLVPPQVRTSYNFGRTFLGLDKC 60  
Db 1 MEPOLGPEAAALRPGWLALLLWVSALSCSFSLSLPASSLSLVPPQVRTSYNFGRTFLGLDKC 60  
Qy 61 NACIGTSICKKFKKEIRSDNWLASHLGLPPDSLSLSPANYSDSKLWRPVEIFRLVSKY 120  
Db 61 NACIGTSICKKFKKEIRSDNWLASHLGLPPDSLSLSPANYSDSKLWRPVEIFRLVSKY 120  
Qy 121 QNEISDRKICASASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOQRELKFLCM 180  
Db 121 QNEISDRKICASASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOQRELKFLCM 180  
Qy 181 LR 182  
Db 181 LR 182

RESULT 14  
US-10-160-503-172  
; Sequence 172, Application US/10160503  
; Publication No. US2004003559A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C446  
; CURRENT APPLICATION NUMBER: US/10/160,503  
; CURRENT FILING DATE: 2002-05-30  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 172  
; LENGTH: 182  
; TYPE: PRT

; ORGANISM: Homo Sapien  
US-10-160-503-172  
Query Match 99.7%; Score 959; DB 12; Length 182;  
Best Local Similarity 99.5%; Pred. No. 1.6e-96;  
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MEPOLGPEAAALRPGWLALLLWVSALSCSFSLSLPASSLSLVPPQVRTSYNFGRTFLGLDKC 60  
Db 1 MEPOLGPEAAALRPGWLALLLWVSALSCSFSLSLPASSLSLVPPQVRTSYNFGRTFLGLDKC 60  
Qy 61 NACIGTSICKKFKKEIRSDNWLASHLGLPPDSLSLSPANYSDSKLWRPVEIFRLVSKY 120  
Db 61 NACIGTSICKKFKKEIRSDNWLASHLGLPPDSLSLSPANYSDSKLWRPVEIFRLVSKY 120  
Qy 121 QNEISDRKICASASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOQRELKFLCM 180  
Db 121 QNEISDRKICASASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOQRELKFLCM 180  
Qy 181 LR 182  
Db 181 LR 182

RESULT 15  
US-10-143-118-172  
; Sequence 172, Application US/10143118  
; Publication No. US20040038335A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C228  
; CURRENT APPLICATION NUMBER: US/10/143,118  
; CURRENT FILING DATE: 2002-05-09  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 172  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-143-118-172

Query Match 99.7%; Score 959; DB 12; Length 182;  
Best Local Similarity 99.5%; Pred. No. 1.6e-96;  
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MEPOLGPEAAALRPGWLALLLWVSALSCSFSLSLPASSLSLVPPQVRTSYNFGRTFLGLDKC 60  
Db 1 MEPOLGPEAAALRPGWLALLLWVSALSCSFSLSLPASSLSLVPPQVRTSYNFGRTFLGLDKC 60  
Qy 61 NACIGTSICKKFKKEIRSDNWLASHLGLPPDSLSLSPANYSDSKLWRPVEIFRLVSKY 120  
Db 61 NACIGTSICKKFKKEIRSDNWLASHLGLPPDSLSLSPANYSDSKLWRPVEIFRLVSKY 120  
Qy 121 QNEISDRKICASASAPKTCSTIERVLRKTERFQKWLQAKRLTPDLVQDCHQOQRELKFLCM 180

Db 121 QNEISDRRICASAPKTCISERVLRKXTERFQKWLQAKRLTPDLVODCHOGQRELLKFLCM 180

QY 181 LR 182

Db 181 LR 182

Search completed: June 8, 2004, 16:31:40  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2004, 16:23:20 ; Search time 22 Seconds  
(without alignments)  
427.088 Million cell updates/sec

Title: US-10-054-988-114  
Perfect score: 962  
Sequence: 1 MEPQLGPEAAALRPGWLALL.....DLVQDCHQGQRELKFLCMLR 182

Scoring table: BLCSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgm2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgm2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgm2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgm2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgm2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgm2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	962	100.0	182	4	US-09-904-615-114
2	962	100.0	209	4	US-09-904-615-168
3	500.5	52.0	146	4	US-09-489-847-161
4	500.5	52.0	146	4	US-09-904-615-121
5	500.5	52.0	146	4	US-09-904-615-169
6	394	41.0	79	4	US-09-904-615-167
7	394	41.0	80	4	US-09-904-615-120
8	85	8.8	420	4	US-09-489-039A-8961
9	80	8.3	896	4	US-09-489-039A-10262
10	79.5	8.3	347	4	US-09-328-352-7107
11	79.5	8.3	928	4	US-09-134-000C-5307
12	76	7.9	271	3	US-09-077-675A-12
13	76	7.9	271	4	US-09-077-674-12
14	76	7.9	302	3	US-09-077-675A-7
15	76	7.9	302	4	US-09-077-674-7
16	76	7.9	361	3	US-09-077-675A-8
17	76	7.9	361	4	US-09-077-674-8
18	76	7.9	366	3	US-09-077-675A-13
19	76	7.9	366	4	US-09-077-674-13
20	76	7.9	366	4	US-09-170-496D-88
21	76	7.9	366	4	US-09-743-742B-7
22	76	7.9	366	4	US-09-762-661A-5
23	76	7.9	366	4	US-09-364-425B-45
24	75.5	7.8	366	4	US-09-743-475-4
25	75.5	7.8	459	4	US-09-252-991A-32058
26	75.5	7.8	1275	3	US-09-120-513-2
27	75.5	7.8	1275	3	US-09-450-105-2

28	75	7.8	302	3	US-09-077-675A-2	Sequence 2, Appli
29	75	7.8	302	4	US-09-077-674-2	Sequence 2, Appli
30	75	7.8	353	3	US-09-077-675A-3	Sequence 3, Appli
31	75	7.8	353	4	US-09-077-674-3	Sequence 3, Appli
32	75	7.8	366	4	US-09-762-661A-7	Sequence 7, Appli
33	75	7.8	366	4	US-09-743-475-6	Sequence 6, Appli
34	74.5	7.7	284	4	US-09-165-827C-14	Sequence 14, Appli
35	74.5	7.7	342	4	US-09-165-827C-2	Sequence 2, Appli
36	74.5	7.7	571	3	US-08-961-083-4	Sequence 4, Appli
37	74.5	7.7	571	4	US-09-536-784-4	Sequence 4, Appli
38	74.5	7.7	1964	2	US-08-790-912-3	Sequence 3, Appli
39	74.5	7.7	2052	2	US-08-790-912-2	Sequence 2, Appli
40	73.5	7.6	1280	4	US-09-672-810-2	Sequence 2, Appli
41	73.5	7.6	1283	4	US-09-672-810-4	Sequence 4, Appli
42	73	7.6	294	4	US-09-424-349A-2	Sequence 2, Appli
43	73	7.6	294	4	US-09-424-349A-9	Sequence 9, Appli
44	72.5	7.5	855	2	US-09-027-337-2	Sequence 2, Appli
45	72.5	7.5	855	4	US-09-644-600-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-904-615-114  
; Sequence 114, Application US/09904615  
; Patent No. 6566325  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 49 Human Secreted Proteins  
; FILE REFERENCE: P2032P1  
; CURRENT APPLICATION NUMBER: US/09/904,615  
; CURRENT FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 09/511,554  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/097,917  
; PRIOR FILING DATE: 1998-08-25  
; PRIOR APPLICATION NUMBER: 60/098,634  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 170  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 114  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-904-615-114

Query Match	100.0%	Score	962	DB	4	Length	182
Best Local Similarity	100.0%	Pred. No.	4	4e-105			
Matches	182	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	MEPQLGPEAAALRPGWLALLMWVSALSCSFSPASSLSLVPGVRTSYNFGRTFLGLDKC	60				
Db	1	MEPQLGPEAAALRPGWLALLMWVSALSCSFSPASSLSLVPGVRTSYNFGRTFLGLDKC	60				
QY	61	NACIGTSICKFKFEKREIRSDNWLASHLGLPPDSLLSYPNKYSDSKIMRPVEIFRLVSKY	120				
Db	61	NACIGTSICKFKFEKREIRSDNWLASHLGLPPDSLLSYPNKYSDSKIMRPVEIFRLVSKY	120				
QY	121	QNEISDRKICASASAPKTCIERVLRKTERFQKWLOAKELTDLVQDCHQGQRELKFLCM	180				
Db	121	QNEISDRKICASASAPKTCIERVLRKTERFQKWLOAKELTDLVQDCHQGQRELKFLCM	180				
QY	181	LR 182					
Db	181	LR 182					

RESULT 2  
US-09-904-615-168  
; Sequence 168, Application US/09904615  
; Patent No. 6566325  
; GENERAL INFORMATION:

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 168
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-904-615-168

Query Match      100.0%; Score 962; DB 4; Length 209;
Best Local Similarity 100.0%; Pred. No. 5.4e-105;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWMLALLWVLSALSCSFSPLPASSLSLVLPQVRTSYNFGRTFLGLDKC 60
DB 28 MEPQLGPEAAALRPGWMLALLWVLSALSCSFSPLPASSLSLVLPQVRTSYNFGRTFLGLDKC 87
QY 61 NACIGTSICKKFFKEIRSDNMLASHLGLPPDSLLSYPNYSDSKIWPRVPEIFRLVSKY 120
DB 88 NACIGTSICKKFFKEIRSDNMLASHLGLPPDSLLSYPNYSDSKIWPRVPEIFRLVSKY 147
QY 121 ONEISDRKICASAPKTCISIEVLKTRFQKWLQAKRLTPDLVODCHOGORELFLCM 180
DB 148 ONEISDRKICASAPKTCISIEVLKTRFQKWLQAKRLTPDLVODCHOGORELFLCM 207
QY 181 LR 182
DB 208 LR 209
```

```
RESULT 3
US-09-489-847-161
; Sequence 161, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 161
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
```

```
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (111)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (122)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (132)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-489-847-161

Query Match      52.0%; Score 500.5; DB 4; Length 146;
Best Local Similarity 72.4%; Pred. No. 6.4e-51;
Matches 105; Conservative 4; Mismatches 33; Indels 3; Gaps 3;

QY 1 MEPQLGPEAAALRPGWMLALLWVLSALSCSFSPLPASSLSLVLPQVRTSYNFGRTFLGLDKC 60
DB 1 MEPQLGPEAAALRPGWMLALLWVLSALSCSFSPLPASSLSLVLPQVRTSYNFGRTFLGLDKC 60
QY 61 NACIGTSICKKFFKEIRSDNMLASHLGLPPDSLLSYPNYSDSKIWPRVPEIFRLVSKY 118
DB 61 NACIGTSICKKFFKEIRSDNMLASHLGLPPDSLLSYPNYSDSKIWPRVPEIFRLVSKY 120
QY 119 KYONEISDRKI-CASASAPKTCISIE 142
DB 121 OXSNEISKOEIXCLLHPKPNLHID 145
```

```
RESULT 4
US-09-904-615-121
; Sequence 121, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 121
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (111)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

; NAME/KEY: SITE  
; LOCATION: (122)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (132)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-904-615-121

Query Match 52.0%; Score 500.5; DB 4; Length 146;  
Best Local Similarity 72.4%; Pred. No. 6.4e-51;  
Matches 105; Conservative 4; Mismatches 33; Indels 3; Gaps 3;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVLPQVTSYNFGRTFLGLDKC 60  
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVLPQVTSYNFGRTFLGLDKC 60  
QY 61 NACIGTSICKKFFKEEIRSDNMLASHLGLPPD-SLSYPANTYSD-DSKIWRPVEIFRLVS 118  
Db 61 NACIGTSICKKFFKEEIRSDNMLASHLGLTASRFLPLXSYFCKLLQMIKIXKXECGXLITGQ 120  
QY 119 KYONEISDRKI-CASASAPKTCISIE 142  
Db 121 QXSNEISKQEIICLLHPPKPNLHID 145

## RESULT 5

US-09-904-615-169  
; Sequence 169, Application US/09904615  
; Patent No. 6566325  
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 49 Human Secreted Proteins  
; FILE REFERENCE: P2032PI  
; CURRENT APPLICATION NUMBER: US/09/904,615  
; PRIOR FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 09/511,554  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/097,917  
; PRIOR FILING DATE: 1998-08-25  
; PRIOR APPLICATION NUMBER: 60/098,634  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 170  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 169  
; LENGTH: 146  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:

; NAME/KEY: SITE  
; LOCATION: (96)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (107)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (111)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (115)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (122)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (132)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-904-615-169

Query Match 52.0%; Score 500.5; DB 4; Length 146;  
Best Local Similarity 72.4%; Pred. No. 6.4e-51;  
Matches 105; Conservative 4; Mismatches 33; Indels 3; Gaps 3;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVLPQVTSYNFGRTFLGLDKC 60

Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVLPQVTSYNFGRTFLGLDKC 60  
QY 61 NACIGTSICKKFFKEEIRSDNMLASHLGLPPD-SLSYPANTYSD-DSKIWRPVEIFRLVS 118  
Db 61 NACIGTSICKKFFKEEIRSDNMLASHLGLTASRFLPLXSYFCKLLQMIKIXKXPGXLITGQ 120  
QY 119 KYONEISDRKI-CASASAPKTCISIE 142  
Db 121 QXSNEISKQEIICLLHPPKPNLHID 145

## RESULT 6

US-09-904-615-167  
; Sequence 167, Application US/09904615  
; Patent No. 6566325  
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 49 Human Secreted Proteins  
; FILE REFERENCE: P2032PI  
; CURRENT APPLICATION NUMBER: US/09/904,615  
; PRIOR FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 09/511,554  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/097,917  
; PRIOR FILING DATE: 1998-08-25  
; PRIOR APPLICATION NUMBER: 60/098,634  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 170  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 167  
; LENGTH: 79  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-904-615-167

Query Match 41.0%; Score 394; DB 4; Length 79;

Best Local Similarity 100.0%; Pred. No. 9.1e-39;  
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVLPQVTSYNFGRTFLGLDKC 60  
Db 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSLPASSLSLVLPQVTSYNFGRTFLGLDKC 60  
QY 61 NACIGTSICKKFFKE 75  
Db 61 NACIGTSICKKFFKE 75

## RESULT 7

US-09-904-615-120  
; Sequence 120, Application US/09904615  
; Patent No. 6566325  
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 49 Human Secreted Proteins  
; FILE REFERENCE: P2032PI  
; CURRENT APPLICATION NUMBER: US/09/904,615  
; PRIOR FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 09/511,554  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/097,917  
; PRIOR FILING DATE: 1998-08-25  
; PRIOR APPLICATION NUMBER: 60/098,634  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 170  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 120  
; LENGTH: 80  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:

; NAME/KEY: SITE



; LOCATION: (80)  
; OTHER INFORMATION: Xaa equals stop translation  
US-09-904-615-120

Query Match 41.0%; Score 394; DB 4; Length 80;  
Best Local Similarity 100.0%; Pred. No. 9.2e-39;  
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEPLQGEAALRPGWALLLWVSALSCSFLPASSLSSLPVQVTSYNGRTFLGLDKC 60  
DB 1 MEPLQGEAALRPGWALLLWVSALSCSFLPASSLSSLPVQVTSYNGRTFLGLDKC 60  
QY 61 NACIGTSICKKFE 75  
DB 61 NACIGTSICKKFE 75

RESULT 8  
US-09-489-039A-8961  
; Sequence 8961, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489.039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 8961  
; LENGTH: 420  
; TYPE: PR1  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-8961

Query Match 8.8%; Score 85; DB 4; Length 420;  
Best Local Similarity 22.2%; Pred. No. 0.22;  
Matches 51; Conservative 31; Mismatches 60; Indels 88; Gaps 13;  
QY 7 PEAAALRPGWALLLWVSALSCSFLPASSLSSLPVQVTSYNGRTFLG 56  
DB 81 PEAAALRPGWALLLWVSALSCSFLPASSLSSLPVQVTSYNGRTFLG 56  
QY 57 LDKC-----NACIGTSIC---KFFKEIRSDNWLASHLGLPDSLLSPANYSDDSKI 107  
DB 123 -DKCPYFYFSDLVVGETTCDGKKWYE---YMAEFKAVHVMQLPNS-----ASDAASRAL 178  
QY 108 WRPEIEIRL---VSKYQNEISDRKI----- 129  
DB 179 WK-TEILRLQOVTEARFGTPISEALREAVLVNRRERLAHFYRLGQLNPPALSGGEIL 237  
QY 130 ----CASAPKTCISIRVLKRTFRF-QKWLQAKRLTPD---LVQDCHQG 171  
DB 238 KVVYGAFFRDKTALIDELHMAERIHQEWQCKLEPRILLITGCPIG 287

RESULT 9  
US-09-489-039A-10262  
; Sequence 10262, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489.039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 10262

; LENGTH: 896  
; TYPE: PR1  
; ORGANISM: Klebsiella pneumoniae  
US-09-039A-10262

Query Match 8.3%; Score 80; DB 4; Length 896;  
Best Local Similarity 23.4%; Pred. No. 2.6; Indels 36; Gaps 6;  
Matches 34; Conservative 18; Mismatches 57; Indels 36; Gaps 6;  
QY 6 GPEAAALRPGWALLLWVSALSCSFLPASSLSSLPVQVTSYNGRTFLGL 57  
DB 269 GPQALAAAGW-----DVAVAAFCGLPADEVAGFYREFIAAPRAITLYTWGINSAGS 323  
QY 58 DKNACIGTSI-----CKKFF-----KKEIRSDNWLASHLGLPDSLLSYP 98  
DB 324 DKNAILINVHLASGKYGRCGPFSLTQCPNMGREYVGLATMLAAHMDVFPDQLRLA 383  
QY 99 ANYSDSKIWPR-----VEIFRLVSK 119  
DB 384 RFWGTERLAQTPGLTAVELFAAIGR 408

RESULT 10  
US-09-328-352-7107  
; Sequence 7107, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328.352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7107  
; LENGTH: 347  
; TYPE: PR1  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7107

Query Match 8.3%; Score 79.5; DB 4; Length 347;  
Best Local Similarity 25.5%; Pred. No. 0.76;  
Matches 25; Conservative 19; Mismatches 37; Indels 17; Gaps 4;  
QY 5 LGP-EAALRPGWALLLWVSALSCSFLPASSLSSLPVQVTSYNGRTFLG 54  
DB 181 LAPLSGVALSPNEILVKILFPNGITLQAVPITNLGNPKKAASVRATWPAIEAQYQFGKS- 239  
QY 55 LGLDKCNACIGTSICKKFF-----KKEIRSDNWLASHL 87  
DB 240 -GVNKFYPLGVGLWYAHFNDIKUNDEIRSLDISAGHM 276

RESULT 11  
US-09-134-000C-5307  
; Sequence 5307, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134.000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 5307  
; LENGTH: 928  
; TYPE: PR1  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-5307

Qy	7	PERAAALPCGWLALLMLWALSALSCSPSLPASSISSIVPOVRSYNFGRFLGLDKNACIGT	66
Db	105	PTEFATRSGLLTVMWVS-SIEFFLPVFCITVLYSLI-----GRKLWRRRGDAVVG	156
Qy	67	SICKPKPKKEIRSDN-----WLASHLG-----LPDSL-----L	95
Db	157	SLRDQNHKKQVTKMLAVVVPALICLMPFFHVGRLFSKFEFGSLIAIQISQYCNLVSFVL	216
Qy	96	SY-----PANYDDSKIWRPVEIFRLVSXYQNEISDRKI	129
Db	217	FYLSAAINPILXIMSKKYR-VAVFRLLG--FFPFSQSKL	253

RESULT 13  
US-09-077-674-12  
\* Sequence 12 Application US/09077674

APPLICANT: Howard, Andrew D.  
APPLICANT: Liberator, Paul A.  
APPLICANT: Schaeffer, James M.  
APPLICANT: Van Der Ploeg, Leonardus  
TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/077,674  
 FILING DATE: 3-JUN-1998  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cocuzzo, Anna L.  
 REGISTRATION NUMBER: 42,452  
 REFERENCE/DOCKET NUMBER: 19589P  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 732-594-1273  
 TELEFAX: 732-594-4720  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 271 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-077-674-12

Db	105	PTEFAVRSGLLTQKQWVS--SIFFLPVFCLTVLYSLI-----GRKLWRRRRGDVAVGA	156
QY	67	SICKKFFKEEIRSDN-----WLASHLG-----LPDSDL-----L	95

Db 157 SERDQNEKQTVKMLAVVVVAFILCWLPFHVGRYLFPSKSPFGSLEIAQISQYCNLVSVL 216  
Qy 96 SY-----PANYSDDSKIMRPVEIFRLVSKYQNEISDRKI 129  
Db 217 FYLSAAINPILYNIMSKKYR-VAVFRLLG--FEFFSQK 253

## RESULT 14

US-09-077-675A-7  
; Sequence 7, Application US/09077675A  
; Patent No. 6242199  
; GENERAL INFORMATION:  
; APPLICANT: Pai, Lee-Yuh  
; APPLICANT: Feighner, Scott C.  
; APPLICANT: Howard, Andrew D.  
; APPLICANT: Pong, Sheng-Shung  
; APPLICANT: Van Der Ploeg, Leonardus H.T.  
; TITLE OF INVENTION: RECEPTOR ASSAY  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merck & Co., Inc.  
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
; CITY: Rahway  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065-0900  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/077,675A  
; FILING DATE: 3-JUN-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cocuzzo, Anna L.  
; REGISTRATION NUMBER: 42,452  
; REFERENCE/DOCKET NUMBER: 19590P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 732-594-1273  
; TELEFAX: 732-594-4720  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 302 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

## US-09-077-675A-7

Query Match 7.9%; Score 76; DB 3; Length 302;  
Best Local Similarity 25.6%; Pred. No. 1.6;  
Matches 41; Conservative 20; Mismatches 51; Indels 48; Gaps 8;  
Qy 7 PEAALRPGLALLLVWVSALSCSFLPASSLSLVPOVRTSYNFGRTFLGLDKCNACIGT 66  
Db 136 PTFEAVRSGLLTVMWVS--SIFFLVPVCLTVLYSLI-----GRKLWRRRGDAVGA 187  
Qy 67 SICKKPFKEIRSDN-----WLASHLG-----LPPDSL-----L 95  
Db 188 SLRDQNEKQTVKMLAVVVVAFILCWLPFHVGRYLFPSKSPFGSLEIAQISQYCNLVSVL 247  
Qy 96 SY-----PANYSDDSKIMRPVEIFRLVSKYQNEISDRKI 129  
Db 248 FYLSAAINPILYNIMSKKYR-VAVFRLLG--FEFFSQK 284

## RESULT 15

US-09-077-674-7  
; Sequence 7, Application US/09077674  
; Patent No. 6531314  
; GENERAL INFORMATION:  
; APPLICANT: Arena, Joseph P.  
; APPLICANT: Cully, Doris F.  
; APPLICANT: Feighner, Scott D.  
; APPLICANT: Howard, Andrew D.  
; APPLICANT: Liberator, Paul A.  
; APPLICANT: Schaeffer, James M.  
; APPLICANT: Van Der Ploeg, Leonardus  
; TITLE OF INVENTION: GROWTH HORMONE SECRETAGOGUE RECEPTOR FAMILY  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merck & Co., Inc.  
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
; CITY: Rahway  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065-0900  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/077,674  
; FILING DATE: 3-JUN-1998  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cocuzzo, Anna L.  
; REGISTRATION NUMBER: 42,452  
; REFERENCE/DOCKET NUMBER: 19589P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 732-594-1273  
; TELEFAX: 732-594-4720  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 302 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

## US-09-077-674-7

Query Match 7.9%; Score 76; DB 4; Length 302;  
Best Local Similarity 25.6%; Pred. No. 1.6;  
Matches 41; Conservative 20; Mismatches 51; Indels 48; Gaps 8;  
Qy 7 PEAALRPGLALLLVWVSALSCSFLPASSLSLVPOVRTSYNFGRTFLGLDKCNACIGT 66  
Db 136 PTFEAVRSGLLTVMWVS--SIFFLVPVCLTVLYSLI-----GRKLWRRRGDAVGA 187  
Qy 67 SICKKPFKEIRSDN-----WLASHLG-----LPPDSL-----L 95  
Db 188 SLRDQNEKQTVKMLAVVVVAFILCWLPFHVGRYLFPSKSPFGSLEIAQISQYCNLVSVL 247  
Qy 96 SY-----PANYSDDSKIMRPVEIFRLVSKYQNEISDRKI 129  
Db 248 FYLSAAINPILYNIMSKKYR-VAVFRLLG--FEFFSQK 284

Search completed: June 8, 2004, 16:26:55

Job time : 23 secs

GenCore version 5.1.6  
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2004, 16:20:14 ; Search time 57 Seconds  
(without alignments)  
902.169 Million cell updates/sec

Title: US-10-054-988-114  
Perfect score: 962  
Sequence: 1 MFPLGPRRALRPGWALL.....DLVQDCEQREKFLCMLR 182

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 1586107 seqs, 282547505 residues  
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	962	100.0	182	3	AAY91393 Human sec
2	962	100.0	182	4	AAB95695 Human pro
3	962	100.0	182	6	ABU03563 Angiogene
4	962	100.0	182	6	ABR47459 Breast ca
5	962	100.0	209	3	AAY91447 Human sec
6	959	99.7	182	4	Aau12257 Human PRO
7	959	99.7	182	4	AAB48066 Human ext
8	959	99.7	182	6	ABO17701 Novel hum
9	959	99.7	182	6	ABU080955 Human PRO
10	959	99.7	182	6	ABU66655 Human PRO
11	959	99.7	182	6	ABU59736 Novel sec
12	959	99.7	182	6	ABO24926 Human sec
13	959	99.7	182	6	ABU66931 Human sec
14	959	99.7	182	6	ADA45691 Novel hum
15	959	99.7	182	6	ADA76122 Human PRO
16	959	99.7	182	6	ADAL8772 Human PRO
17	959	99.7	182	6	ADA61395 Homo sapi
18	959	99.7	182	6	ADBI19180 Novel hum
19	959	99.7	182	6	ADB27721 Human PRO
20	959	99.7	182	6	ADA86200 Novel hum
21	959	99.7	182	6	ADBI5764 Human PRO
22	959	99.7	182	6	ADA47550 Human PRO
23	959	99.7	182	6	ADA67345 Human PRO
24	959	99.7	182	6	ADB30352 Human PRO
25	959	99.7	182	6	ADA85648 Novel hum

## ALIGNMENTS

RESULT 1  
AAY91393  
ID AAY91393 standard; protein; 182 AA.  
XX  
AC AAY91393;  
XX  
DT 29-JUN-2000 (first entry)  
XX  
DE Human secreted protein sequence encoded by gene 48 SEQ ID NO:114.  
XX  
KW Human; secreted protein; diagnosis; neuroprotective; nootropic;  
KW neuroleptic; antinaptic; cerebroprotective; immunomodulatory;  
KW anti-microbial; cardiant; cytostatic; antiinflammatory; haemostatic;  
KW anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy;  
KW neural; reproductive; immune disorder; immunodeficiency; infection;  
KW lymphoma; demyelinating disease; autoimmunity; cancer; inflammation;  
KW neyruism; haemorrhage; Alzheimer's disease; Parkinson's disease;  
KW Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis;  
KW ischaemia; mania; dementia; obsessive compulsive disorder;  
KW viral prophylaxis; developmental disorder; sexually-linked disorder;  
KW cardiovascular disorder; food additive; preservative.  
XX  
OS Homo sapiens.  
XX  
PN WO200011014-A1.  
XX  
PD 02-MAR-2000.  
XX  
PF 24-AUG-1999; 99WO-US019330.  
XX  
PR 25-AUG-1998; 98US-0097917P.  
PR 31-AUG-1998; 98US-0098634P.  
XX  
(HUMA-) HUMAN GENOME SCI INC.  
PI Moore PA, Ruben SM, Olsen HS, Shi Y, Rosen CA, Florence KA;  
PI Soppet DR, Lafleur DW, Endress GA, Ebner R, Komatsoulis G, Duan RD;  
XX WPI; 2000-224656/19.  
DR N-PSDB; AAA26328.  
XX  
XX Novel secreted proteins and corresponding DNA molecules that can be used  
PT to prevent, treat and diagnose disease in humans, for example,  
PT Alzheimer's, cancer, and immune disorders.  
XX  
XX Claim 11; Page 380-381; 416pp; English.  
XX  
CC The polynucleotide sequences given in AAA26281 to AAA26336 encode the

26 959 99.7 182 6 ADA96860  
27 959 99.7 182 6 ADA79164  
28 959 99.7 182 6 ADA87303  
29 959 99.7 182 6 ADB16505  
30 959 99.7 182 6 ADA91597  
31 959 99.7 182 6 ADB14660  
32 959 99.7 182 6 ADB18621  
33 959 99.7 182 6 ADA93836  
34 959 99.7 182 6 ADB19732  
35 959 99.7 182 6 ADB13044  
36 959 99.7 182 6 ABO43234  
37 959 99.7 182 6 ADA74298  
38 959 99.7 182 6 ADB24531  
39 959 99.7 182 6 ADA82055  
40 959 99.7 182 6 ADA75018  
41 959 99.7 182 6 ADA85096  
42 959 99.7 182 6 ADA84544  
43 959 99.7 182 6 ADB29800  
44 959 99.7 182 6 ADA80328  
45 959 99.7 182 6 ADA75570

Ada96860 Human PRO  
Ada79164 Human PRO  
Ada87303 Novel hum  
Adb16505 Human PRO  
Ada91597 Novel hum  
Adb14660 Human PRO  
Adb18621 Novel hum  
Ada93836 Human PRO  
Adb19732 Novel hum  
Adb13044 Human PRO  
Abo43234 Novel hum  
Ada74298 Human PRO  
Adb24531 Human PRO  
Ada82055 Human PRO  
Ada75018 Human PRO  
Ada85096 Novel hum  
Ada84544 Novel hum  
Adb29800 Human PRO  
Ada80328 Human PRO  
Ada75570 Human PRO

CC human secreted proteins given in AAY91346 to AAY91449. The human secreted  
 CC proteins can have activities based on the tissues and cells they are  
 CC expressed in. Examples of the activities are: neuroprotective; nootropic;  
 CC neuroleptic; antiamic; cerebroprotective; immunomodulatory; anti-  
 CC microbial; cardiant; cytostatic; antiinflammatory; haemostatic;  
 CC anticonvulsant; and vasotrophic. The polynucleotides and proteins may be  
 CC used to prevent, treat or ameliorate a medical condition, e.g. by protein  
 CC or gene therapy. Conditions treatable by the proteins of the invention  
 CC include neural, reproductive, or immune disorders, especially  
 CC immunodeficiency, infection, lymphomas, demyelinating diseases, auto-  
 CC immunities, cancer, general microbial infection, inflammation, aneurysms  
 CC and haemorrhages. Specific examples include: Alzheimer's disease;  
 CC Parkinson's; Huntington's; Tourette syndrome; multiple sclerosis;  
 CC meningitis; ischaemia; prostate cancer; mania; dementia; obsessive  
 CC compulsive disorder and viral prophylaxis. The polynucleotides and  
 CC proteins can also be used in the detection of disorders associated with  
 CC the function of the protein, for example, the detection of developmental  
 CC disorders, sexually-linked disorders, or disorders of the cardiovascular  
 CC system. They may also be used as food additives or preservatives.  
 CC AAA26272 to AAA26280 and AAY91345 are sequences used in the  
 CC exemplification of the present invention

XX SQ Sequence 182 AA;

Query Match 100.0%; Score 962; DB 3; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-95;  
 Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFLSPASSLSLVLPQVRTSYNFGRTFLGLDKC 60  
 |||||  
 DB 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFLSPASSLSLVLPQVRTSYNFGRTFLGLDKC 60  
 |||||  
 QY 61 NACIGTSICKKFFKEIRSDNWLASHLGLPDSLLSYPNYSDSKTWRRPVEIFRLVSKY 120  
 |||||  
 DB 61 NACIGTSICKKFFKEIRSDNWLASHLGLPDSLLSYPNYSDSKTWRRPVEIFRLVSKY 120  
 |||||  
 QY 121 ONEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQGRELKFLCM 180  
 |||||  
 DB 121 ONEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQGRELKFLCM 180  
 |||||  
 QY 181 LR 182  
 ||  
 DB 181 LR 182

RESULT 2

AA95695  
 ID AAB95695 standard; protein; 182 AA.

XX AAB95695;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:18516.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX PA

PI Oza T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
 PT length cDNAs defined in the specification, and for the detection and/or  
 PT diagnosis of the abnormality of the proteins encoded by the full-length  
 PT cDNAs.

PS Claim 8; SEQ ID NO 18516; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesising 5602 full-  
 CC length cDNAs defined in the specification. Where a primer set comprises:  
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
 CC complementary strand of a polynucleotide which comprises one of the 5602  
 CC nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
 CC specification. The primer sets can be used in antisense therapy and in  
 CC gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AA92446 to AA95893  
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
 CC oligonucleotides, all of which are used in the exemplification of the  
 CC present invention

XX SQ Sequence 182 AA;

Query Match 100.0%; Score 962; DB 4; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-95;  
 Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFLSPASSLSLVLPQVRTSYNFGRTFLGLDKC 60  
 |||||  
 DB 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFLSPASSLSLVLPQVRTSYNFGRTFLGLDKC 60  
 |||||  
 QY 61 NACIGTSICKKFFKEIRSDNWLASHLGLPDSLLSYPNYSDSKTWRRPVEIFRLVSKY 120  
 |||||  
 DB 61 NACIGTSICKKFFKEIRSDNWLASHLGLPDSLLSYPNYSDSKTWRRPVEIFRLVSKY 120  
 |||||  
 QY 121 ONEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQGRELKFLCM 180  
 |||||  
 DB 121 ONEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQGRELKFLCM 180  
 |||||  
 QY 181 LR 182  
 ||  
 DB 181 LR 182

RESULT 3

ABU03563

ID ABU03563 standard; protein; 182 AA.

XX ABU03563;

XX 21-JAN-2003 (first entry)

XX Angiogenesis-associated human protein sequence #108.

XX Human; angiogenesis-associated transcript; angiogenesis;

XX angiogenesis-associated disease; cancer; cytostatic.

XX Homo sapiens.

XX OS

```

PN WC200279492-A2.
XX
PD 10-OCT-2002.
XX
PF 14-FEB-2002; 2002WO-US004915.
XX
PR 14-FEB-2001; 2001US-00784356.
XX
PR 22-FEB-2001; 2001US-00791390.
PR 19-APR-2001; 2001US-0285475P.
PR 03-AUG-2001; 2001US-0310025P.
PR 13-NOV-2001; 2001US-0350866P.
PR 29-NOV-2001; 2001US-0334244P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Murray R, Glynn R, Watson SR, Aziz N;
PI
XX
XX WPI; 2003-040681/03.
XX N-PSDB; ABX08947.
XX
PT Detecting angiogenesis-associated transcript in a cell for diagnosing and
PT treating cancer by contacting a sample with a polynucleotide that
PT exhibits changes in expression level as a function of time in tissue
PT undergoing angiogenesis.
XX
XX Example 2; Page 282; 291pp; English.
XX
XX The present invention relates to methods and compositions for detecting
XX an angiogenesis-associated transcript in a cell in a patient. The method
XX involves contacting a biological sample from the patient with a
XX polynucleotide that selectively hybridizes to a sequence at least 80%
XX identical to any of the angiogenesis-associated human polynucleotide
XX sequences given in the specification. These angiogenesis-associated
XX polynucleotide sequences comprise genes that exhibit changes in
XX expression levels as a function of time in tissue undergoing
XX angiogenesis. The method and the polynucleotide sequences of the
XX invention are useful for diagnosing and treating angiogenesis and
XX angiogenesis-associated diseases e.g. cancer. The polynucleotide
XX sequences are also useful in the gene therapy of such disorders. The
XX angiogenesis-associated proteins encoded by the polynucleotide sequences
XX are useful as a vaccine for therapeutic and prophylactic immunisation.
XX ABJ03456-ABJ03569 represent angiogenesis-associated protein sequences
XX
XX Sequence 182 AA;
XX
XX Query Match 100.0%; Score 962; DB 6; Length 182;
XX Best Local Similarity 100.0%; Pred. No. 3.6e-95;
XX Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MEPQGPAAALRPGWMLALLWVSALSCSFSLPASSLSLVQVRTSYNFGRTFLGLDKC 60
Db |||||
QY 1 MEPQGPAAALRPGWMLALLWVSALSCSFSLPASSLSLVQVRTSYNFGRTFLGLDKC 60
Db |||||
QY 61 NACIGTSICKKFFKEIRSDNWLASHGLPPDSLSLYPANYSDSKINRPVIFRLVSKY 120
Db |||||
QY 61 NACIGTSICKKFFKEIRSDNWLASHGLPPDSLSLYPANYSDSKINRPVIFRLVSKY 120
Db |||||
QY 121 ONEISDRKICASAPKTCSTIERVLKTERFQKWLQAKRLTPDLVQDCHQOGRELKFLCM 180
Db |||||
QY 121 ONEISDRKICASAPKTCSTIERVLKTERFQKWLQAKRLTPDLVQDCHQOGRELKFLCM 180
Db |||||
QY 181 LR 182
Db |||||
QY 181 LR 182
Db |||||
XX
XX RESULT 4
XX ABR47459
XX ID ABR47459 standard; protein; 182 AA.
XX
XX AC ABR47459;
XX
XX DT 12-JUN-2003 (first entry)

```

```

XX
DE Breast cancer associated protein sequence SEQ ID NO:150.
XX
XX Human; breast cancer; cytostatic; gene therapy.
XX
XX Homo sapiens.
XX
XX WO2003004989-A2.
XX
XX 16-JAN-2003.
XX
XX 21-JUN-2002; 2002WO-US019669.
XX
XX 21-JUN-2001; 2001US-0299887P.
XX 27-JUN-2001; 2001US-0301572P.
XX 18-JUL-2001; 2001US-0306501P.
XX 25-SEP-2001; 2001US-0325002P.
XX 05-MAR-2002; 2002US-0362585P.
XX 14-MAY-2002; 2002US-0380391P.
XX
XX (MILL-) MILLENIUM PHARM INC.
XX
XX Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;
XX Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
XX Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB,
XX
XX WPI; 2003-210381/20.
XX N-PSDB; ACC50151.
XX
XX Breast cancer diagnosis or treatment by comparing the level of expression
XX of a marker in a patient sample with that in the control non-breast
XX cancer sample.
XX
XX Claim 1; SEQ ID NO 150; 128pp; English.
XX
XX The present invention describes a method for assessing whether a patient
XX is afflicted with breast cancer. The method comprises comparing the level
XX of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
XX ABR47386 to ABR47632) in a patient sample and the normal level of
XX expression of the marker in a control non-breast cancer sample, where a
XX significant increase in the level of expression of the marker in the
XX patient sample and the normal level is an indication that the patient
XX is afflicted with breast cancer. The breast cancer associated sequences from
XX the present invention have cytostatic activities and can be used in gene
XX therapy. The method is useful for diagnosing and treating breast cancer.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 182 AA;
XX
XX Query Match 100.0%; Score 962; DB 6; Length 182;
XX Best Local Similarity 100.0%; Pred. No. 3.6e-95;
XX Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MEPQGPAAALRPGWMLALLWVSALSCSFSLPASSLSLVQVRTSYNFGRTFLGLDKC 60
Db |||||
QY 1 MEPQGPAAALRPGWMLALLWVSALSCSFSLPASSLSLVQVRTSYNFGRTFLGLDKC 60
Db |||||
QY 61 NACIGTSICKKFFKEIRSDNWLASHGLPPDSLSLYPANYSDSKINRPVIFRLVSKY 120
Db |||||
QY 61 NACIGTSICKKFFKEIRSDNWLASHGLPPDSLSLYPANYSDSKINRPVIFRLVSKY 120
Db |||||
QY 121 ONEISDRKICASAPKTCSTIERVLKTERFQKWLQAKRLTPDLVQDCHQOGRELKFLCM 180
Db |||||
QY 121 ONEISDRKICASAPKTCSTIERVLKTERFQKWLQAKRLTPDLVQDCHQOGRELKFLCM 180
Db |||||
QY 181 LR 182
Db |||||
QY 181 LR 182
Db |||||
XX
XX RESULT 5

```

AA99:447  
 ID AAY91447 standard; protein; 209 AA.  
 AC AAY91447;  
 XX  
 DT 29-JUN-2000 (first entry)  
 DE  
 XX Human secreted protein sequence encoded by gene 48 SEQ ID NO:168.  
 KW Human; secreted protein; diagnosis; neuroprotective; nootropic;  
 KW neuroleptic; antimanic; cerebroprotective; immunomodulatory;  
 KW anti-microbial; cardiant; cytostatic; antiinflammatory; haemostatic;  
 KW anticonvulsant; vasotropic; vaccine; gene therapy; anti-sense therapy;  
 KW neural; reproductive; immune disorder; immunodeficiency; infection;  
 KW lymphoma; demyelinating disease; autoimmune; cancer; inflammation;  
 KW aneurysm; haemorrhage; Alzheimer's disease; Parkinson's disease;  
 KW Huntington's disease; Tourette syndrome; multiple sclerosis; meningitis;  
 KW ischaemia; mania; dementia; obsessive compulsive disorder;  
 KW viral prophylaxis; developmental disorder; sexually-linked disorder;  
 KW cardiovascular disorder; food additive; preservative.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200011014-A1.  
 XX  
 XX 02-MAR-2000.  
 XX  
 XX 24-AUG-1999; 99WO-US019330.  
 XX  
 XX 25-AUG-1998; 98US-0097917P.  
 XX 31-AUG-1998; 98US-0098634P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Moore PA, Ruben SM, Olsen HS, Shi Y, Rosen CA, Florence KA;  
 XX Soppet DR, Lafleur DW, Endress GA, Ebner R, Komatsoulis G, Duan RD;  
 XX WPI; 2000-224656/19.  
 XX  
 XX Novel secreted proteins and corresponding DNA molecules that can be used  
 XX to prevent, treat and diagnose disease in humans, for example,  
 XX Alzheimer's, cancer, and immune disorders.  
 XX  
 XX Disclosure; Page 413-414; 416pp; English.  
 XX  
 XX The polynucleotide sequences given in AAA26281 to AAA26336 encode the  
 XX human secreted proteins given in AAY91346 to AAY91449. The human secreted  
 XX proteins can have activities based on the tissues and cells they are  
 XX expressed in. Examples of the activities are: neuroprotective; nootropic;  
 XX neuroleptic; antimanic; cerebroprotective; immunomodulatory; anti-  
 XX microbial; cardiant; cytostatic; antiinflammatory; haemostatic;  
 XX anticonvulsant; and vasotropic. The polynucleotides and proteins may be  
 XX used to prevent, treat or ameliorate a medical condition, e.g. by protein  
 XX or gene therapy. Conditions treatable by the proteins of the invention  
 XX include neural, reproductive, or immune disorders, especially  
 XX immunodeficiency, infection, lymphomas, demyelinating diseases, auto-  
 XX immunities, cancer, general microbial infection, inflammation, aneurysms  
 XX and haemorrhages. Specific examples include: Alzheimer's disease;  
 XX Parkinson's; Huntington's; Tourette syndrome; multiple sclerosis;  
 XX meningitis; ischaemia; prostate cancer; mania; dementia; obsessive  
 XX compulsive disorder and viral prophylaxis. The polynucleotides and  
 XX proteins can also be used in the detection of disorders associated with  
 XX the function of the protein, for example, the detection of developmental  
 XX disorders, sexually-linked disorders, or disorders of the cardiovascular  
 XX system. They may also be used as food additives or preservatives.  
 XX AAA26272 to AAA26280 and AAY91345 are sequences used in the  
 XX exemplification of the present invention  
 XX  
 XX Sequence 209 AA;

Query Match 100.0%; Score 962; DB 3; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-95;  
 Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLLALLLWVSALSCSFSI PASSLSLVQVTSYNGRTFLGLDKC 60  
 DB |||||  
 QY 28 MEPQLGPEAAALRPGWLLALLLWVSALSCSFSI PASSLSLVQVTSYNGRTFLGLDKC 87  
 DB |||||  
 QY 61 NACIGTSICKKFFKEIRSDNWLASHLGLPPDSLLSYDPANYSDDSKIWPRVFI FRLVSKY 120  
 DB |||||  
 QY 88 NACIGTSICKKFFKEIRSDNWLASHLGLPPDSLLSYDPANYSDDSKIWPRVFI FRLVSKY 147  
 DB |||||  
 QY 121 QNEISDRKICASAPKTCSTIERVLKTRFKWLQAKRLTPDLVQDCHQGORELKFLCM 180  
 DB |||||  
 QY 148 QNEISDRKICASAPKTCSTIERVLKTRFKWLQAKRLTPDLVQDCHQGORELKFLCM 207  
 DB |||||  
 QY 181 LR 182  
 DB ||  
 QY 208 LR 209  
 DB ||  
 RESULT 6  
 AAU12257  
 ID AAU12257 standard; protein; 182 AA.  
 XX  
 AC AAU12257;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human PRO3743 polypeptide sequence.  
 XX  
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;  
 KW prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;  
 KW ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;  
 KW A-peptide; factor VIIA; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200140466-A2.  
 XX  
 XX 07-JUN-2001.  
 XX  
 XX 01-DEC-2000; 2000WO-US032678.  
 XX  
 PR 01-DEC-1999; 99WO-US028301.  
 PR 01-DEC-1999; 99WO-US028634.  
 PR 02-DEC-1999; 99WO-US028551.  
 PR 02-DEC-1999; 99WO-US028564.  
 PR 02-DEC-1999; 99WO-US028565.  
 PR 09-DEC-1999; 99US-0170262P.  
 PR 16-DEC-1999; 99WO-US030095.  
 PR 20-DEC-1999; 99WO-US030911.  
 PR 20-DEC-1999; 99WO-US030999.  
 PR 30-DEC-1999; 99WO-US031243.  
 PR 30-DEC-1999; 99WO-US031274.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000277.  
 PR 06-JAN-2000; 2000WO-US0003176.  
 PR 11-FEB-2000; 2000WO-US0003565.  
 PR 18-FEB-2000; 2000WO-US000341.  
 PR 18-FEB-2000; 2000WO-US0004342.  
 PR 22-FEB-2000; 2000WO-US000414.  
 PR 24-FEB-2000; 2000WO-US0004914.  
 PR 24-FEB-2000; 2000WO-US005004.  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 02-MAR-2000; 2000WO-US005641.  
 PR 03-MAR-2000; 2000US-0187202P.  
 PR 10-MAR-2000; 2000WO-US006319.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 20-MAR-2000; 2000WO-US007377.  
 PR 21-MAR-2000; 2000WO-US007532.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.

```

PR 05-JUN-2000; 2000US-02098322.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
XX (GETH ) GENENTECH INC.
PI Baker XZ, Bersini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
DR WPI: 2001-408281/43.
DR N-PSDB; AAS21329.
XX
XX Isolated , secretory and transmembrane PRO polypeptide used to detect
XX other PRO polypeptides, link bioactive molecules to cells expressing PRO
XX polypeptides, and detect the presence of mammalian tumors e.g. lung,
XX breast, prostate, cervical.
XX
XX Claim 12; Fig 172; 813pp; English.
XX
XX AAU12172-AAU12446 represent novel human secretory and transmembrane PRO
XX polypeptides. The PRO polypeptides are useful to detect other PRO
XX polypeptides, to link bioactive molecules to cells expressing PRO
XX polypeptides, to modulate biological activities of cells expressing PRO
XX polypeptides, and to detect the presence of mammalian lung, colon,
XX breast, prostate, rectal, cervical or liver tumours by comparing PRO
XX polypeptide expression in a cell sample to that in a control sample. Some
XX of the 275 sequences are also useful to stimulate the release of tumour
XX necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or
XX differentiation of chondrocytes, the proliferation or gene expression in
XX pericyte cells, the release of proteoglycans from cartilage, the
XX proliferation of inner ear utricular supporting cells or of T-
XX lymphocytes, the release of a cytokine from peripheral blood monocytes
XX (PBMCs), or the proliferation of endothelial cells. Some of the PRO
XX polypeptides may modulate glucose or free fatty acid uptake by skeletal
XX muscle cells or by adipocytes; or inhibit binding of A-peptide to factor
XX VIIA. The PRO polypeptides can be used in assays to identify molecules
XX involved in binding interactions. The polynucleotides encoding PRO
XX polypeptides can be used to generate probes, antisense RNA/DNA,
XX transgenic or knock out animals and can be used in gene therapy
XX
XX Sequence 182 AA;
XX
XX Query Match          99.7%; Score 959; DB 4; Length 182;
XX Best Local Similarity 99.5%; Pred. No. 7.6e-95;
XX Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MEPQLGPRAALRPGWALLLLWVSLSCFSLPASSLSLVLPQVRTSYNFGRTFLGLKRC 60
Db 1 MEPQLGPRAALRPGWALLLLWVSLSCFSLPASSLSLVLPQVRTSYNFGRTFLGLKRC 60
QY 61 NACIGTSICKFFKEIEISDNWLASHGLPPDLSISYPANYSDDSKIMRPVIEIRLWSKY 120
Db 61 NACIGTSICKFFKEIEISDNWLASHGLPPDLSISYPANYSDDSKIMRPVIEIRLWSKY 120
QY 121 QNEISDRKTCASAPKTCISIEVLKRTFRFQKWLQAKELTDLVDQCHQGQRELKFLCM 180
Db 121 QNEISDRKTCASAPKTCISIEVLKRTFRFQKWLQAKELTDLVDQCHQGQRELKFLCM 180
QY 181 LR 182
Db 181 LR 182

```

```

RESULT 7
AAB48066
ID AAB48066 standard; protein; 182 AA.
XX
XX AAB48066;
AC

```

```

XX 19-MAR-2001 (first entry)
XX Human extracellular signaling molecule (EXCS) (ID 5090841CD1).
XX
XX Extracellular signaling molecule; EXCS; anti-inflammatory; human;
XX immunosuppressive; cytostatic; neuroprotective; gastrointestinal;
XX virucide; antibacterial; anti-HIV; human immunodeficiency virus;
XX antiinfectility; cerebroprotective; nootropic; antitumor; antifungal;
XX anticonvulsant; tranquilizer; neuroleptic; vasotropic; gynecological;
XX keratolytic; protozoacide; gene therapy.
XX Homo sapiens.
XX
XX WO200070049-A2.
XX
XX 23-NOV-2000.
XX
XX 19-MAY-2000; 2000WO-US013975.
XX
XX 19-MAY-1999; 99US-0134949P.
XX 15-JUL-1999; 99US-0144270P.
XX 30-JUL-1999; 99US-0146700P.
XX 04-OCT-1999; 99US-0157508P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Tang YT, Yue H, Lal P, Burford N, Bandman O, Baughn MR;
XX Azimzai Y, Lu DM, Patterson C;
XX
XX WPI: 2001-025021/03.
XX N-PSDB; AAC84302.
XX
XX New human extracellular signaling nucleic acids and polypeptides useful
XX for diagnosing, treating and preventing infections and gastrointestinal,
XX neurological, reproductive, and autoimmune/inflammatory disorders.
XX
XX Claim 1; Page 88-89; 114pp; English.
XX
XX The invention provides human extracellular signaling molecules (EXCS) and
XX polynucleotides which identify and encode EXCS. EXCS can be expressed by
XX standard recombinant methodology. The amino acid and nucleic acid
XX sequences of EXCS are useful for diagnosing, treating and preventing
XX infections and gastrointestinal (peptic ulcer, dysphagia, pancreatitis),
XX neurological (e.g. epilepsy, ischemic cerebrovascular disease, stroke),
XX reproductive (infertility, ovulatory defects, endometriosis), autoimmune
XX /inflammatory (actinic keratosis, acquired immunodeficiency syndrome
XX (AIDS), Addison's disease), and cell proliferative disorders including
XX cancers (of the breast, adrenal gland, bone). They may also be used to
XX treat fatal familial insomnia, nutritional and metabolic diseases of the
XX nervous system, myopathies, mental disorders (anxiety, schizophrenia,
XX mood), as well as infections caused by parasites (malaria, leishmania,
XX trypanosoma), viral (adenovirus, coronavirus, flavivirus), bacterial
XX (e.g. pneumococcus, staphylococcus, bacillus), and fungal (aspergillus,
XX blastomycetes, dermatophytes) agents. The nucleic acids, polypeptides,
XX antagonists, agonists, pharmaceutical compositions, and antibodies may
XX also be used for treating or preventing disorders associated with
XX increased or decreased expression or activity of EXCS. EXCS
XX polynucleotides may also be used to detect and quantify gene expression
XX in biopsied tissues in which expression of EXCS may be correlated with
XX the disease, to determine presence or excess expression of EXCS, to
XX monitor regulation of EXCS levels during therapeutic intervention, to
XX detect the presence of associated disorders, as targets in microarray, to
XX generate hybridization probes, and to detect differences in gene
XX sequences among normal, carrier or affected individuals. Antibodies may
XX also be used in diagnosing disorders, in monitoring patients being
XX treated with EXCS agonists, antagonists or inhibitors. Sequences AAB48057
XX -B48082 represent the EXCS of the invention
XX
XX Sequence 182 AA;
XX
XX Query Match          99.7%; Score 959; DB 4; Length 182;
XX Best Local Similarity 99.5%; Pred. No. 7.6e-95;

```



Page 6

PR 19-DEC-2001; 2001US-00028072.  
XX (GETH) GENENTECH INC.  
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX WPI: 2003-341980/32.  
DR N-PSDB; ACD23338.  
XX  
XX New secreted and transmembrane PRO nucleic acids, for treating  
PT inflammation, organ failure, atherosclerosis, cardiac injury,  
PT infertility, birth defects, premature aging, acquired immunodeficiency  
XX syndrome (AIDS), or cancer.  
XX  
XX Claim 12; Fig 172; 660pp; English.  
XX The invention describes an isolated nucleic acid (I) comprising, or which  
CC has 80 % sequence identity to, or the full-length coding sequence of, one  
CC of 275 nucleotide sequences, and which encodes a corresponding  
CC polypeptide selected from 275 amino acid sequences, where all sequences  
CC are given in the specification. The polypeptide encoded by (I) is used to  
CC detect PRO polypeptides, link a bioactive molecule to a cell, stimulate the  
CC PRO polypeptide, modulate a biological activity of a cell, stimulate the  
CC release of tumor necrosis factor (TNF)-alpha from human blood, modulate  
CC the uptake of glucose or free fatty acid by cells, stimulate or inhibit  
CC the proliferation or differentiation of cells or gene expression,  
CC stimulate the release of proteoglycans, stimulate the release of cytokine  
CC from peripheral blood mononuclear cells, inhibit the binding of A-peptide  
CC to factor VIIA, or detect the presence of tumor in a mammal. The nucleic  
CC acid and polypeptide encoded by it, are useful for treating inflammatory  
CC diseases, organ failure, atherosclerosis, cardiac injury, infertility,  
CC birth defects, premature aging, acquired immunodeficiency syndrome  
CC (AIDS), cancer, or diabetic complications. The nucleic acid is useful as  
CC hybridisation probes, in chromosome and gene mapping, and in generating  
CC antisense RNA or DNA. The polypeptides are useful as pharmaceuticals,  
CC diagnostics, biosensors or bioreactors. Both are useful in tissue typing.  
CC This is the amino acid sequence of a novel human secreted and  
CC transmembrane PRO polypeptide  
XX  
XX Sequence 182 AA;  
SQ  
Query Match 99.7%; Score 959; DB 6; Length 182;  
Best Local Similarity 99.5%; Pred. No. 7.6e-95;  
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEPQLGPEAAALRPGMLALLVSVLSLSCSFLPSSLSILVQVRTSYNFGRTFLGLDKC 60  
DB 1 MEPQLGPEAAALRPGMLALLVSVLSLSCSFLPSSLSILVQVRTSYNFGRTFLGLDKC 60  
QY 61 NACITSTICKKFKKEIRSDNWLASHLGLPPDSLSYPANYSDDSKIMRPVIFRLVSKY 120  
DB 61 NACITSTICKKFKKEIRSDNWLASHLGLPPDSLSYPANYSDDSKIMRPVIFRLVSKY 120  
QY 121 ONEISDRKTCASAPKTCISIRVLKRTKTFKQKLOAKLTLPDLVDDCHOGORELKFLCM 180  
DB 121 ONEISDRKTCASAPKTCISIRVLKRTKTFKQKLOAKLTLPDLVDDCHOGORELKFLCM 180  
QY 181 LR 182  
DB 181 LR 182  
RESULT 9  
ID ABU0355  
XX ABU0955 standard; protein; 182 AA.  
XX AC ABU0955;  
XX  
XX 23-JUN-2003 (first entry)  
XX  
XX Human PRO polypeptide #86.

XX Human; PRO polypeptide; secreted and transmembrane protein;  
XX anti-PRO antibody; diagnostic assay; gene expression; diabetes;  
XX bone disorder; cartilage disorder; rheumatoid arthritis; obesity;  
XX sports injury; osteoarthritis; hyper-insulinaemia; hypo-insulinaemia;  
XX hearing loss; coagulation disorder; stroke; heart attack; cardiac;  
XX antidiabetic; anorectic; vulnery; antiarthritis; osteopathic;  
XX antirheumatic; auditory; cerebroprotective; angiogenic.  
XX  
XX Homo sapiens.  
XX US20030043111-A1.  
XX  
XX 02-JAN-2003.  
XX  
XX 19-DEC-2001; 2001US-00028072.  
XX  
XX 18-JUN-1997; 97US-0049911P.  
XX 26-AUG-1997; 97US-0056974P.  
XX 17-SEP-1997; 97US-0059113P.  
XX 17-SEP-1997; 97US-0059115P.  
XX 17-SEP-1997; 97US-0059117P.  
XX 17-SEP-1997; 97US-0059122P.  
XX 18-SEP-1997; 97US-0059184P.  
XX 19-SEP-1997; 97US-0059263P.  
XX 19-SEP-1997; 97US-0059352P.  
XX 24-SEP-1997; 97US-0059588P.  
XX 17-OCT-1997; 97US-0062250P.  
XX 17-OCT-1997; 97US-0062285P.  
XX 17-OCT-1997; 97US-0062287P.  
XX 17-OCT-1997; 97US-0063755P.  
XX 24-OCT-1997; 97US-0063814P.  
XX 24-OCT-1997; 97US-0063816P.  
XX 24-OCT-1997; 97US-0063045P.  
XX 24-OCT-1997; 97US-0063082P.  
XX 27-OCT-1997; 97US-0063127P.  
XX 27-OCT-1997; 97US-0063327P.  
XX 27-OCT-1997; 97US-0063329P.  
XX 28-OCT-1997; 97US-0063505P.  
XX 29-OCT-1997; 97US-0063561P.  
XX 29-OCT-1997; 97US-0063704P.  
XX 29-OCT-1997; 97US-0063733P.  
XX 29-OCT-1997; 97US-0063735P.  
XX 29-OCT-1997; 97US-0063738P.  
XX 03-NOV-1997; 97US-0064248P.  
XX 07-NOV-1997; 97US-0064809P.  
XX 12-NOV-1997; 97US-0065186P.  
XX 17-NOV-1997; 97US-0065846P.  
XX 21-NOV-1997; 97US-0066364P.  
XX 24-NOV-1997; 97US-0066453P.  
XX 24-NOV-1997; 97US-0066511P.  
XX 24-NOV-1997; 97US-0066770P.  
XX 11-DEC-1997; 97US-0069212P.  
XX 11-DEC-1997; 97US-0069278P.  
XX 11-DEC-1997; 97US-0069334P.  
XX 16-DEC-1997; 97US-0069694P.  
XX 23-JAN-1998; 98US-0072320P.  
XX 04-FEB-1998; 98US-0073612P.  
XX 09-FEB-1998; 98US-0074086P.  
XX 09-FEB-1998; 98US-0074092P.  
XX 12-MAR-1998; 98US-0077791P.  
XX 20-MAR-1998; 98US-0078910P.  
XX 25-MAR-1998; 98US-0079294P.  
XX 27-MAR-1998; 98US-0079663P.  
XX 27-MAR-1998; 98US-0079728P.  
XX 31-MAR-1998; 98US-0080165P.  
XX 12-JUN-1998; 98WO-US012456.  
XX 14-JUL-1998; 98WO-US014552.  
XX 28-AUG-1998; 98WO-US017888.  
XX 10-SEP-1998; 98WO-US018824.  
XX 14-SEP-1998; 98WO-US019093.  
XX 14-SEP-1998; 98WO-US019094.

No PRO 3743



14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 98WO-US000106.  
PR 08-MAR-1999; 98WO-US005028.  
PR 10-MAR-1999; 98WO-US005190.  
PR 20-APR-1999; 98WO-US008615.  
PR 14-MAY-1999; 98WO-US010733.  
PR 02-JUN-1999; 98WO-US012252.  
PR 01-SEP-1999; 98WO-US020111.  
PR 08-SEP-1999; 98WO-US020594.  
PR 13-SEP-1999; 98WO-US020944.  
PR 15-SEP-1999; 98WO-US021090.  
PR 05-OCT-1999; 98WO-US021547.  
PR 05-OCT-1999; 98WO-US023089.  
PR 29-NOV-1999; 98WO-US028214.  
PR 30-NOV-1999; 98WO-US028313.  
PR 30-NOV-1999; 98WO-US028409.  
PR 01-DEC-1999; 98WO-US028301.  
PR 01-DEC-1999; 98WO-US028634.  
PR 02-DEC-1999; 98WO-US028551.  
PR 02-DEC-1999; 98WO-US028564.  
PR 02-DEC-1999; 98WO-US028565.  
PR 16-DEC-1999; 98WO-US030095.  
PR 20-DEC-1999; 98WO-US030911.  
PR 20-DEC-1999; 98WO-US030999.  
PR 30-DEC-1999; 98WO-US031243.  
PR 30-DEC-1999; 98WO-US031274.  
PR 05-JAN-2000; 2000WO-US030219.  
PR 06-JAN-2000; 2000WO-US030277.  
PR 06-JAN-2000; 2000WO-US030376.  
PR 11-FEB-2000; 2000WO-US030365.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
XX XX  
(GETH ) GENENTECH INC.

Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen MB, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX  
WPI; 2003-352836/33.  
DR N-PSDB; ACA67079.  
XX  
PT New isolated PRO polypeptide useful for treating diabetes, rheumatoid  
PT arthritis, sports injuries, obesity, hearing loss in mammals, stroke, or  
XX heart attack.  
XX

Claim 12; Fig 172; 643pp; English.

XX  
CC The present invention relates to the isolation of novel human PRO  
CC polypeptides, and the polynucleotide sequences encoding them. The PRO  
CC polypeptides are secreted and transmembrane proteins. The PRO  
CC polypeptides and polynucleotides are useful for preparing a medicament  
CC useful in the treatment of diabetes, bone and/or cartilage disorders  
CC (e.g. rheumatoid arthritis, sports injuries, osteoarthritis), obesity,  
CC hyper- or hypo-insulinaemia, hearing loss, and coagulation disorders  
CC (e.g. stroke, heart attack). Anti-PRO antibodies are useful in diagnostic  
CC assays for PRO, by detecting its expression in specific cells, tissues or  
CC serum, and for affinity purification of PRO from recombinant cell culture  
CC or natural sources. AB08070-AB081144 represent the human PRO  
CC polypeptides of the invention. Note: The sequence data for this patent  
CC was obtained in electronic format directly from the USPTO web site at

CC seqdata.uspto.gov/psipidsIDEntry.html  
XX  
SQ Sequence 182 AA;  
Query Match 99.7%; Score 959; DB 6; Length 182;  
Best Local Similarity 99.5%; Pred. No. 7.6e-95;  
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSIPASSLSLVLPQVTSYNFGRFTFLGLDKC 60  
DB 1 MEPQLGPEAAALRPGWLALLLWVSALSCSFSIPASSLSLVLPQVTSYNFGRFTFLGLDKC 60  
QY 61 NACIGTSICKKPFKEETRSNDNLASHLGLPDSLSLYPANTYSDDSKIWRPVEIFPLVSKY 120  
DB 61 NACIGTSICKKPFKEETRSNDNLASHLGLPDSLSLYPANTYSDDSKIWRPVEIFPLVSKY 120  
QY 121 QNEISDRKICASAPKTCSTIERYLRTKTERFQKWLQAKRLTPDLVQDCHQORELKLCLM 180  
DB 121 QNEISDRKICASAPKTCSTIERYLRTKTERFQKWLQAKRLTPDLVQDCHQORELKLCLM 180  
QY 181 LR 182  
DB 181 LR 182  
RESULT 10  
ABU66655  
ID ABU66655 standard; protein; 182 AA.  
XX AC ABU66655;  
XX DT 23-MAY-2003 (first entry)  
XX DE Human PRO polypeptide #86.  
XX KW Human; PRO polypeptide; secreted and transmembrane protein;  
XX KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;  
XX KW differentiation; chondrocyte; tumour; genetic disorder; cytostatic.  
XX OS Homo sapiens.  
XX PN US2003036180-A1.  
XX PD 20-FEB-2003.  
XX PF 09-MAY-2002; 2002US-00143114.  
XX PR 31-MAR-1997; 97WO-US005230.  
XX PR 12-JUN-1998; 98WO-US012456.  
XX PR 14-JUL-1998; 98WO-US014552.  
XX PR 28-AUG-1998; 98WO-US017888.  
XX PR 10-SEP-1998; 98WO-US018824.  
XX PR 14-SEP-1998; 98WO-US019093.  
XX PR 14-SEP-1998; 98WO-US019094.  
XX PR 14-SEP-1998; 98WO-US019177.  
XX PR 15-SEP-1998; 98WO-US019330.  
XX PR 17-SEP-1998; 98WO-US019437.  
XX PR 07-OCT-1998; 98WO-US021141.  
XX PR 29-OCT-1998; 98WO-US022991.  
XX PR 20-NOV-1998; 98WO-US022992.  
XX PR 01-DEC-1998; 98WO-US024855.  
XX PR 05-JAN-1999; 98WO-US025108.  
XX PR 08-MAR-1999; 98WO-US005028.  
XX PR 10-MAR-1999; 98WO-US005190.  
XX PR 20-APR-1999; 98WO-US008615.  
XX PR 14-MAY-1999; 98WO-US010733.  
XX PR 02-JUN-1999; 98WO-US012252.  
XX PR 01-SEP-1999; 98WO-US020111.  
XX PR 08-SEP-1999; 98WO-US020594.  
XX PR 13-SEP-1999; 98WO-US020944.  
XX PR 15-SEP-1999; 98WO-US021090.  
XX PR 15-SEP-1999; 98WO-US021547.

```
PR 05-OCT-1999; 99WO-US023089.
PR 29-NOV-1999; 99WO-US028214.
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US003376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 21-MAR-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US034956.
PR 20-DEC-2000; 2000WO-US034959.
PR 28-FEB-2001; 2001WO-US003649.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001WO-US008270.
PR 14-MAR-2001; 2001WO-US008689.
PR 22-MAR-2001; 2001WO-US008174.
PR 05-APR-2001; 2001WO-US008236.
PR 10-MAY-2001; 2001WO-US008542.
PR 10-MAY-2001; 2001WO-US008542.
PR 18-MAY-2001; 2001WO-US008602.
PR 25-MAY-2001; 2001WO-US008602.
PR 25-MAY-2001; 2001WO-US008603.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001WO-US008745.
PR 14-JUN-2001; 2001WO-US008263.
PR 19-JUN-2001; 2001WO-US008834.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001WO-US008879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001WO-US009088.
PR 06-AUG-2001; 2001WO-US009241.
PR 09-AUG-2001; 2001WO-US009279.
PR 16-AUG-2001; 2001WO-US009318.
PR 19-DEC-2001; 2001WO-US002802.

XX (GETH ) GENENTECH INC.
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
XX Gerritsen ME, Goddard A, Godowski PU, Gurney AL, Sherwood S;
XX Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2003-332040/31.
XX N-PSDB; ACA03688.
XX New secreted and transmembrane PRO nucleic acids, useful for gene
XX therapy, in chromosome and gene mapping, as chromosome markers, in tissue
XX typing, and in chromosome identification.
XX Claim 12; Fig 172; 660pp; English.
XX The present invention relates to the isolation of novel human PRO
XX polypeptides, and the polynucleotide sequences encoding them. The PRO
XX polypeptides are secreted and transmembrane proteins. The PRO
XX polypeptides are useful for detecting other PRO polypeptides, for linking
XX bioactive molecules to cells expressing PRO polypeptides, for modulating
XX biological activities of cells expressing PRO polypeptides, and for
XX identifying agonists or antagonists. The PRO polypeptides are useful for
XX for stimulating the release of tumour necrosis factor (TNF)-alpha from
XX human blood, for stimulating the proliferation or differentiation of
XX chondrocytes, and detecting the presence of tumours. The polynucleotide
XX sequences encoding PRO polypeptides are useful as hybridisation probes,
XX in chromosome and gene mapping, in the generation of antisense RNA and
XX DNA, in the preparation of PRO polypeptides, for generating transgenic
XX animals or knockout animals, for the genetic analysis of individuals with
XX genetic disorders, and in gene therapy. ABU66570-ABU66844 represent the
XX human PRO polypeptides of the invention. Note: The sequence data for this
XX patent was obtained in electronic format directly from the USPTO web site
XX at seqdata.uspto.gov/psipsdEntry.html
XX Sequence 182 AA;

Query Match 99.7%; Score 959; DB 6; Length 182;
Best Local Similarity 99.5%; Pred. NO. 7.6e-95;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALPGWLALLWVSALSCSFSIPASSLSILVPOVRTSYNFGRTFLGLDKC 60
Db 1 MEPQLGPEAAALPGWLALLWVSALSCSFSIPASSLSILVPOVRTSYNFGRTFLGLDKC 60
QY 61 NACIGTSICKKFFKEIRSDNWLASHGLPPDLSILSYPNYSDSKIMRPVEIFRLVSKY 120
Db 61 NACIGTSICKKFFKEIRSDNWLASHGLPPDLSILSYPNYSDSKIMRPVEIFRLVSKY 120
QY 121 ONEISDRKICASAPKTCISIERVLKTRFQKWLQAKELTDLVDQCHQGQRELKFLCM 180
Db 121 ONEISDRKICASAPKTCISIERVLKTRFQKWLQAKELTDLVDQCHQGQRELKFLCM 180
QY 181 LR 182
Db 181 LR 182

RESULT 11
ABU59736
ID ABU59736 standard; protein; 182 AA.
XX AC ABU59736;
XX DT 13-MAY-2003 (first entry)
XX DE Novel secreted and transmembrane protein PRO3743.
XX Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
KW cardiac insufficiency disorder; cancer; tumour; immune response;
KW adrenal cortical capillary endothelial growth; c-fos induction;
KW vascular endothelial growth factor inhibition; VEGF inhibition;
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
```

retinal neurons cell survival; rod photoreceptor cell survival;  
 retinal disorder; retinitis pigmentosa; kidney disorder;  
 mammalian kidney mesangial cell proliferation; Berger disease;  
 dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;  
 chondrocyte redifferentiation; sports injury; arthritis.  
 Homo sapiens.  
 OS  
 XX  
 XX  
 US2003017563-A1.  
 XX  
 PD  
 23-JAN-2003.  
 XX  
 XX  
 PF  
 07-MAY-2002; 2002US-00140808.  
 XX  
 PR  
 31-MAR-1997; 97WO-US005230.  
 PR  
 12-JUN-1998; 98WO-US012466.  
 PR  
 14-JUL-1998; 98WO-US014552.  
 PR  
 28-AUG-1998; 98WO-US017888.  
 PR  
 10-SEP-1998; 98WO-US018824.  
 PR  
 14-SEP-1998; 98WO-US019093.  
 PR  
 14-SEP-1998; 98WO-US019094.  
 PR  
 14-SEP-1998; 98WO-US019177.  
 PR  
 16-SEP-1998; 98WO-US019330.  
 PR  
 17-SEP-1998; 98WO-US019437.  
 PR  
 07-OCT-1998; 98WO-US021141.  
 PR  
 29-OCT-1998; 98WO-US022991.  
 PR  
 20-NOV-1998; 98WO-US024855.  
 PR  
 01-DEC-1998; 98WO-US025108.  
 PR  
 05-JAN-1999; 99WO-US000106.  
 PR  
 08-MAR-1999; 99WO-US005028.  
 PR  
 10-MAR-1999; 99WO-US005190.  
 PR  
 20-APR-1999; 99WO-US008615.  
 PR  
 14-MAY-1999; 99WO-US010733.  
 PR  
 02-JUN-1999; 99WO-US012252.  
 PR  
 01-SEP-1999; 99WO-US020111.  
 PR  
 08-SEP-1999; 99WO-US020594.  
 PR  
 13-SEP-1999; 99WO-US020944.  
 PR  
 15-SEP-1999; 99WO-US021090.  
 PR  
 15-SEP-1999; 99WO-US021547.  
 PR  
 05-OCT-1999; 99WO-US023089.  
 PR  
 29-NOV-1999; 99WO-US028214.  
 PR  
 30-NOV-1999; 99WO-US028313.  
 PR  
 30-NOV-1999; 99WO-US028409.  
 PR  
 01-DEC-1999; 99WO-US028301.  
 PR  
 01-DEC-1999; 99WO-US028634.  
 PR  
 02-DEC-1999; 99WO-US028551.  
 PR  
 02-DEC-1999; 99WO-US028564.  
 PR  
 02-DEC-1999; 99WO-US028565.  
 PR  
 16-DEC-1999; 99WO-US030095.  
 PR  
 20-DEC-1999; 99WO-US030911.  
 PR  
 20-DEC-1999; 99WO-US030999.  
 PR  
 22-DEC-1999; 99WO-US030720.  
 PR  
 30-DEC-1999; 99WO-US031243.  
 PR  
 30-DEC-1999; 99WO-US031274.  
 PR  
 05-JAN-2000; 2000WO-US000219.  
 PR  
 06-JAN-2000; 2000WO-US000277.  
 PR  
 11-FEB-2000; 2000WO-US003365.  
 PR  
 18-FEB-2000; 2000WO-US004341.  
 PR  
 18-FEB-2000; 2000WO-US004342.  
 PR  
 22-FEB-2000; 2000WO-US004414.  
 PR  
 24-FEB-2000; 2000WO-US004914.  
 PR  
 24-FEB-2000; 2000WO-US005004.  
 PR  
 01-MAR-2000; 2000WO-US005601.  
 PR  
 02-MAR-2000; 2000WO-US005746.  
 PR  
 02-MAR-2000; 2000WO-US005841.  
 PR  
 10-MAR-2000; 2000WO-US006319.  
 PR  
 15-MAR-2000; 2000WO-US006884.  
 PR  
 20-MAR-2000; 2000WO-US007377.  
 PR  
 21-MAR-2000; 2000WO-US007532.  
 PR  
 30-MAR-2000; 2000WO-US008439.  
 PR  
 17-MAY-2000; 2000WO-US013705.  
 PR  
 22-MAY-2000; 2000WO-US014042.  
 PR  
 30-MAY-2000; 2000WO-US014941.  
 PR  
 02-JUN-2000; 2000WO-US015264.  
 PR  
 28-JUL-2000; 2000WO-US020710.  
 PR  
 11-AUG-2000; 2000WO-US022031.  
 PR  
 23-AUG-2000; 2000WO-US023522.  
 PR  
 24-AUG-2000; 2000WO-US023328.  
 PR  
 08-NOV-2000; 2000WO-US030952.  
 PR  
 10-NOV-2000; 2000WO-US030873.  
 PR  
 01-DEC-2000; 2000WO-US032678.  
 PR  
 20-DEC-2000; 2000US-00747259.  
 PR  
 20-DEC-2000; 2000WO-US034956.  
 PR  
 28-FEB-2001; 2001US-00796498.  
 PR  
 28-FEB-2001; 2001WO-US006520.  
 PR  
 01-MAR-2001; 2001WO-US006666.  
 PR  
 09-MAR-2001; 2001US-00802706.  
 PR  
 14-MAR-2001; 2001US-00808689.  
 PR  
 22-MAR-2001; 2001US-00816744.  
 PR  
 05-APR-2001; 2001US-00828366.  
 PR  
 10-MAY-2001; 2001US-00854208.  
 PR  
 18-MAY-2001; 2001US-00860216.  
 PR  
 25-MAY-2001; 2001US-00866028.  
 PR  
 25-MAY-2001; 2001US-00866034.  
 PR  
 25-MAY-2001; 2001US-00867092.  
 PR  
 01-JUN-2001; 2001US-00872035.  
 PR  
 01-JUN-2001; 2001WO-US017800.  
 PR  
 05-JUN-2001; 2001US-00874503.  
 PR  
 14-JUN-2001; 2001US-00882636.  
 PR  
 19-JUN-2001; 2001US-00886342.  
 PR  
 20-JUN-2001; 2001WO-US019692.  
 PR  
 21-JUN-2001; 2001US-00887879.  
 PR  
 22-JUN-2001; 2001WO-US020116.  
 PR  
 29-JUN-2001; 2001WO-US021066.  
 PR  
 09-JUL-2001; 2001WO-US021735.  
 PR  
 18-JUL-2001; 2001US-00908927.  
 PR  
 06-AUG-2001; 2001US-00924419.  
 PR  
 09-AUG-2001; 2001US-00927796.  
 PR  
 16-AUG-2001; 2001US-00931836.  
 PR  
 19-DEC-2001; 2001US-00028072.  
 XX  
 PA  
 (GETH ) GENENTECH INC.  
 XX  
 XX  
 PI  
 Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 PI  
 Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI  
 Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX  
 WPI; 2003-148238/14.  
 N-PSDB; ABX89226.  
 XX  
 Two hundred and seventy five nucleic acids encoding PRO polypeptides,  
 useful for treating pericyte-associated tumors, diabetes and various bone  
 and/or cartilage disorders, e.g. arthritis.  
 Claim 12; Fig 172; 659pp; English.  
 XX  
 The invention describes an isolated human PRO polypeptide. The PRO  
 polypeptides are useful in detecting PRO polypeptides in a sample, in  
 linking a bioactive molecule to a cell expressing a PRO polypeptide, and  
 in modulating at least one biological activity of a cell expressing a PRO  
 polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus  
 useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186  
 stimulate adrenal cortical capillary endothelial growth, and PRO536,  
 PRO943, PRO828, PRO1068 or PRO535, PRO826, PRO819, PRO1126,  
 PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus  
 useful for treating conditions or disorders where angiogenesis would be  
 beneficial, e.g. wound healing and antagonist of this polypeptide are  
 useful for treating cancerous tumours. PRO812 inhibits vascular  
 endothelial growth factor (VEGF) stimulated proliferation of endothelial  
 cells and is thus useful for inhibiting endothelial cell growth in  
 mammals which would be beneficial in inhibiting tumour growth. PRO826,  
 PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of  
 stimulated T-lymphocytes and are therapeutically useful for enhancing

CC immune response. PRO828, PRO825, PRO1068 or PRO1132 enhance survival of  
CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of  
CC rod photoreceptor cells) and therefore are useful for treating retinal  
CC disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813  
CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,  
CC and therefore are useful for treating kidney disorders associated with  
CC decreased mesangial cell function such as Berger disease or other  
CC nephropathies associated with dermatitis, herpetiformis or Crohn's  
CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the  
CC proliferation and/or redifferentiation of chondrocytes in culture and are  
CC thus useful for treating sports injuries, and arthritis. This is the  
CC amino acid sequence of a novel human PRO protein  
XX  
SQ Sequence 182 AA;

Query Match 99.7%; Score 959; DB 6; Length 182;

Best Local Similarity 99.5%; Pred. No. 7.6e-95;

Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPLGPEAAALRPGWLALLLVSLSCSFSPASSLSILVPQVTSYNFGTELGLOK 60  
DB 1 MEPLGPEAAALRPGWLALLLVSLSCSFSPASSLSILVPQVTSYNFGTELGLOK 60  
QY 61 NACIGTSICKKFKKEIRSDNWLASHGLPPDSLLSYPNYSDDSKIWPFVEIFRLVSKY 120  
DB 61 NACIGTSICKKFKKEIRSDNWLASHGLPPDSLLSYPNYSDDSKIWPFVEIFRLVSKY 120  
QY 121 QNEISDRKTCASAPKTCISIERVLKTERFQKWLQAKELTDLVDCHQGOELKPLCM 180  
DB 121 QNEISDRKTCASAPKTCISIERVLKTERFQKWLQAKELTDLVDCHQGOELKPLCM 180  
QY 181 LR 182  
DB 181 LR 182

## RESULT 12

ABO24926

ID ABO24926 standard; protein; 182 AA.

XX AC ABO24926;

XX DT 05-SEP-2003 (first entry)

XX DE Human secreted/transmembrane protein (PRO) #86.

XX KW Human; PRO; secreted protein; transmembrane protein; tumour; cytostatic;  
XX KW gene therapy; tumour necrosis factor-alpha; TNF-alpha; blood;  
XX KW proteoglycan; cartilage; cytokine; peripheral blood mononuclear cell;  
XX KW PBMC; glucose uptake; PFA; skeletal muscle cell; adipocyte cell;  
XX KW chondrocyte cell proliferation; chondrocyte cell differentiation;  
XX KW pericyte cell; inner ear utricular supporting cell; T-lymphocyte cell;  
XX KW endothelial cell; A-peptide; factor VIIa.

XX CS Homo sapiens.

XX OS US2003036179-A1.

XX PN 20-FEB-2003.

XX PF 10-MAY-2002; 2002US-00142431.

XX PR 31-MAR-1997; 97WO-US005230.

XX PR 12-JUN-1998; 98WO-US012456.

XX PR 14-JUL-1998; 98WO-US014552.

XX PR 28-AUG-1998; 98WO-US017888.

XX PR 10-SEP-1998; 98WO-US018824.

XX PR 14-SEP-1998; 98WO-US019093.

XX PR 14-SEP-1998; 98WO-US019094.

XX PR 16-SEP-1998; 98WO-US019177.

XX PR 17-SEP-1998; 98WO-US019330.

XX PR 17-SEP-1998; 98WO-US019437.

XX PR 07-OCT-1998; 98WO-US021141.

XX PR 31-MAR-1997; 97WO-US005230.

XX PR 12-JUN-1998; 98WO-US012456.

XX PR 14-JUL-1998; 98WO-US014552.

XX PR 28-AUG-1998; 98WO-US017888.

XX PR 10-SEP-1998; 98WO-US018824.

XX PR 14-SEP-1998; 98WO-US019093.

XX PR 14-SEP-1998; 98WO-US019094.

XX PR 16-SEP-1998; 98WO-US019177.

XX PR 17-SEP-1998; 98WO-US019330.

XX PR 17-SEP-1998; 98WO-US019437.

XX PR 07-OCT-1998; 98WO-US021141.

PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 98WO-US000106.  
PR 08-MAR-1999; 98WO-US005028.  
PR 10-MAR-1999; 98WO-US005190.  
PR 20-APR-1999; 98WO-US008615.  
PR 14-MAY-1999; 98WO-US010733.  
PR 02-JUN-1999; 98WO-US012252.  
PR 01-SEP-1999; 98WO-US020111.  
PR 08-SEP-1999; 98WO-US020594.  
PR 13-SEP-1999; 98WO-US020944.  
PR 15-SEP-1999; 98WO-US021090.  
PR 15-SEP-1999; 98WO-US021547.  
PR 05-OCT-1999; 98WO-US023089.  
PR 29-NOV-1999; 98WO-US028214.  
PR 30-NOV-1999; 98WO-US028313.  
PR 30-NOV-1999; 98WO-US028409.  
PR 01-DEC-1999; 98WO-US028301.  
PR 01-DEC-1999; 98WO-US028634.  
PR 02-DEC-1999; 98WO-US028551.  
PR 02-DEC-1999; 98WO-US028564.  
PR 16-DEC-1999; 98WO-US030095.  
PR 20-DEC-1999; 98WO-US030911.  
PR 20-DEC-1999; 98WO-US030999.  
PR 22-DEC-1999; 98WO-US030720.  
PR 30-DEC-1999; 98WO-US031243.  
PR 30-DEC-1999; 98WO-US031274.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000277.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004342.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 01-MAR-2000; 2000WO-US005601.  
PR 02-MAR-2000; 2000WO-US005746.  
PR 10-MAR-2000; 2000WO-US005841.  
PR 15-MAR-2000; 2000WO-US006319.  
PR 20-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 21-MAR-2000; 2000WO-US007532.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 28-FEB-2001; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 10-MAY-2001; 2001US-00854280.  
PR 18-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001WO-US017092.

PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001WO-US021666.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tamas D, Watanabe CK, Wood WL, Zhang Z;  
XX  
DR WPI; 2003-466355/44.  
DR N-PSDB; ACD41880.  
XX

XX New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO1114 or  
PT PRO4978, useful in molecular biology, chromosome and gene mapping, in  
PT generating antisense RNA and DNA, and in gene therapy.  
XX

XX Claim 12; Fig 172; 659pp; English.

XX The invention relates to an isolated nucleic acid comprising at least 80t  
CC sequence identity to a PRO (secreted and transmembrane protein) cDNA  
CC comprising a nucleic acid (a) encoding a PRO polypeptide, or its  
CC extracellular domain (with or without its associated signal peptide),  
CC which comprises any of the 275 120-850 residue amino acid sequences,  
CC given in the specification; (b) comprising any of the 275 300-3500'  
CC nucleotide sequences, given in the specification; or (c) comprising the  
CC full-length coding sequence of the nucleotide sequences given in the  
CC specification, or of the DNA deposited under any of the American Type  
CC Culture Collection (ATCC) Accession Numbers listed in the specification.  
CC Also included are a vector comprising the novel nucleic acid, a host cell  
CC comprising the vector, producing a PRO polypeptide, the isolated PRO  
CC polypeptides detailed above, a chimeric molecule comprising the PRO  
CC polypeptide of fused to a heterologous amino acid sequence, an anti-PRO  
CC antibody, detecting a PRO polypeptide in a sample suspected of containing  
CC the PRO polypeptide, linking a bioactive molecule to a cell expressing a  
CC PRO polypeptide, modulating at least one biological activity of a cell  
CC expressing a PRO polypeptide, stimulating the release of tumour necrosis  
CC factor-alpha (TNF-alpha) from human blood, (or proteoglycans from  
CC cartilage or cytokine from peripheral blood mononuclear cells (PBMC)),  
CC modulating the uptake of glucose or FFA by skeletal muscle cells or  
CC adipocyte cells, stimulating the proliferation or differentiation of  
CC chondrocyte cells (for proliferation or gene expression in pericyte  
CC cells), stimulating the proliferation of inner ear utricular supporting  
CC cells (or of T-lymphocyte cells, or of endothelial cells), inhibiting the  
CC binding of A-peptide to factor VIIA, or differentiation of adipocyte  
CC cells, detecting the presence of a tumour in a mammal and an  
CC oligonucleotide probe derived from any of the nucleotide sequences given  
CC in the specification. The polynucleotide is useful in molecular biology,  
CC including uses as hybridisation probes, in chromosome and gene mapping,  
CC in generating antisense RNA and DNA, and in gene therapy. The  
CC polynucleotide may also be used in preparing PRO polypeptides by  
CC recombinant techniques, and in generating either transgenic animals or  
CC knock-out animals which, in turn, are useful in the development and  
CC screening of therapeutically useful reagents. The PRO polypeptide or the  
CC antibody is used in preparing a medicament for treating a condition  
CC responsive to the polypeptide or antibody, such as tumours, and in  
CC various diagnostic assays. The present sequence represents a PRO  
CC polypeptide

XX Sequence 182 AA;  
SQ

Query Match 99.7%; Score 959; DB 6; Length 182;  
Best Local Similarity 99.5%; Pred. No. 7.6e-95;  
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEPQLGPEAAALRPGWLALLLWVSLSCFSLSLSPASSLSLSPVQVRTSYNFGRTFLGLDKC 60  
DB 1 MEPQLGPEAAALRPGWLALLLWVSLSCFSLSLSPASSLSLSPVQVRTSYNFGRTFLGLDKC 60  
QY 61 NACIGTSICKKFFKEIRSDNWLASHLGLPDPDLSLLSPANYSDDSKIWRPVEIFRLVSKY 120  
DB 61 NACIGTSICKKFFKEIRSDNWLASHLGLPDPDLSLLSPANYSDDSKIWRPVEIFRLVSKY 120  
QY 121 ONEISDRKICASAPKTCSTIERVLRKTERFKMLQAKRLTPDLVQDCHQORELKLFCM 180  
DB 121 ONEISDRKICASAPKTCSTIERVLRKTERFKMLQAKRLTPDLVQDCHQORELKLFCM 180  
QY 181 LR 182  
DB 181 LR 182  
RESULT 13  
ABU66931  
ID ABU66931 standard; protein; 182 AA.  
XX AC ABU66931;  
XX DT 27-MAY-2003 (first entry)  
XX DE Human secreted/transmembrane, PRO, protein SEQ ID 172.  
XX KW Human; secreted protein; transmembrane protein; PRO;  
KW inflammatory disease; organ failure; atherosclerosis; cardiac injury;  
KW infertility; birth defects; premature aging; AIDS; biosensor;  
KW acquired immunodeficiency syndrome; cancer; diabetic complication;  
XX bioreactor; tumour.  
XX OS Homo sapiens.  
XX PN US2003032155-A1.  
XX PD 13-FEB-2003.  
XX PP 03-MAY-2002; 2002US-00137865.  
PR 31-MAR-1997; 97WO-US005230.  
PR 12-JUN-1998; 98WO-US012456.  
PR 14-JUL-1998; 98WO-US014552.  
PR 28-AUG-1998; 98WO-US017888.  
PR 10-SEP-1998; 98WO-US018824.  
PR 14-SEP-1998; 98WO-US019093.  
PR 14-SEP-1998; 98WO-US019094.  
PR 14-SEP-1998; 98WO-US019177.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 29-OCT-1998; 98WO-US022991.  
PR 29-OCT-1998; 98WO-US022992.  
PR 20-NOV-1998; 98WO-US024855.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 10-MAR-1999; 99WO-US005190.  
PR 20-APR-1999; 99WO-US008615.  
PR 14-MAY-1999; 99WO-US010733.  
PR 02-JUN-1999; 99WO-US012252.  
PR 01-SEP-1999; 99WO-US020111.  
PR 08-SEP-1999; 99WO-US020594.  
PR 13-SEP-1999; 99WO-US020944.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 05-OCT-1999; 99WO-US023089.  
PR 29-NOV-1999; 99WO-US028214.

```
PR 30-NOV-1999; 99WO-US028313.
PR 30-NOV-1999; 99WO-US028409.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028564.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030395.
PR 20-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030999.
PR 22-DEC-1999; 99WO-US030720.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 18-FEB-2000; 2000WO-US004342.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005746.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 10-NOV-2000; 2000WO-US030873.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000US-0034956.
PR 28-FEB-2001; 2001US-00796498.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 09-MAR-2001; 2001US-00802706.
PR 14-MAR-2001; 2001US-00808689.
PR 22-MAR-2001; 2001US-00816744.
PR 05-APR-2001; 2001US-00828366.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 18-MAY-2001; 2001US-00860216.
PR 25-MAY-2001; 2001US-00866028.
PR 25-MAY-2001; 2001US-00866034.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 21-JUN-2001; 2001US-00897879.
PR 22-JUN-2001; 2001WO-US020116.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 18-JUL-2001; 2001US-00908827.
PR 06-AUG-2001; 2001US-00924419.
PR 09-AUG-2001; 2001US-00927796.
PR 16-AUG-2001; 2001US-00931836.
PR 19-DEC-2001; 2001US-00028072.
XX (GETH ) GENENTECH INC.

XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski RJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX N-PSDB; ACA04109.
WPI; 2003-331925/31.
DR N-PSDB; ACA04109.
XX
PT New secreted and transmembrane nucleic acids and polypeptides, designated
PT as PRO, useful for treating inflammation, organ failure, atherosclerosis,
PT cardiac injury, infertility, birth defects, premature aging, AIDS, or
PT cancer.
XX
PS Claim 12; Fig 172; 659pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising, or which is
CC at least 80% identical to, or the full-length coding sequence of, any of
CC the 275 nucleotide sequences, encoding the corresponding PRO polypeptide
CC (one of 275 secreted or transmembrane proteins). The nucleic acid further
CC comprises the full-length coding sequence of the DNA deposited under
CC American Type Culture Collection (ATCC) accession number in a list given
CC in the specification. Also included are vectors and host cells for
CC producing PRO proteins, PRO fusion proteins, anti-PRO antibodies, PRO
CC extracellular domains and mature sequences, methods of detecting PRO
CC proteins, methods for stimulating the release of TNF-alpha (tumour
CC necrosis factor alpha) from human blood, (and the proliferation of
CC differentiation of chondrocyte cells, the proliferation of, or gene
CC expression in pericyte cells, the release of proteoglycans from
CC cartilage, proliferation of inner ear uricular supporting cells, the
CC proliferation of T-lymphocyte cells, the release of a cytokine from
CC peripheral blood mononuclear cells (PBMC), or the proliferation of
CC endothelial cells), a method for modulating the uptake of glucose or free
CC fatty acid (FFA) by skeletal muscle cells, a method for inhibiting the
CC binding of A-peptide to factor VIIa, or the differentiation of adipocyte
CC cells, a method for detecting the presence of a tumour in a mammal and an
CC oligonucleotide probe derived from any of the nucleotide sequences cited
CC above. The nucleic acids and polypeptides are useful for treating
CC inflammatory diseases, organ failure, atherosclerosis, cardiac injury,
CC infertility, birth defects, premature aging, AIDS (acquired
CC immunodeficiency syndrome), cancer, or diabetic complications. The
CC nucleic acids are useful as hybridisation probes, in chromosome and gene
CC mapping, and in generating antisense RNA or DNA. The polypeptides are
CC useful as pharmaceuticals, diagnostics, biosensors or bioeffectors. Both
CC are useful in tissue typing. The present sequence represents a PRO
CC protein of the invention
XX
SQ Sequence 182 AA;
Query Match 99.7%; Score 959; DB 6; Length 182;
Best Local Similarity 99.5%; Pred. No. 7.6e-95;
Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEPQLGPEAAALRPGWMLALLWVSALSCFSFSLPSSLSLVPOVRTSYNFGRTFLGLDKC 60
DB 1 MEPQLGPEAAALRPGWMLALLWVSALSCFSFSLPSSLSLVPOVRTSYNFGRTFLGLDKC 60
QY 61 NACIGTSICKPKFKKEIRSDNWLASHGLPPDLSLSYPANYSOSKIWPVEIFRLVSKY 120
DB 61 NACIGTSICKPKFKKEIRSDNWLASHGLPPDLSLSYPANYSOSKIWPVEIFRLVSKY 120
QY 121 QNEISDRKICASAPKTCISIERVLKTERFKWLQAKRLTDLVDQCHQGRELKFLCM 180
DB 121 QNEISDRKICASAPKTCISIERVLKTERFKWLQAKRLTDLVDQCHQGRELKFLCM 180
QY 181 LR 182
DB 181 LR 182
RESULT 14
ADA45691
ID ADA45691 standard; protein; 182 AA.
XX
```





CC for stimulating the proliferation of T-lymphocyte cells, for stimulating  
 CC the release of a cytokine from PBMC cells, for inhibiting the binding of  
 CC A-peptide to factor VIIa, for inhibiting the differentiation of adipocyte  
 CC cells, for stimulating proliferation of endothelial cells, for detecting  
 CC the presence of tumour in a mammal. The tumour is lung, colon, breast,  
 CC prostate, rectal, cervical or liver tumour. The oligonucleotide probes  
 CC are useful for isolating genomic and cDNA nucleotide sequences or  
 CC antisense probes. (I) is also useful as therapeutic agent. PRO is useful  
 CC in assays to identify other proteins or molecules involved in binding  
 CC interaction. A polynucleotide (II) encoding (I) is useful in chromosome  
 CC and gene mapping, in generation of antisense RNA and DNA, in the  
 CC preparation of PRO polypeptide, for generating transgenic animals or  
 CC knockout animals which in turn are useful in the development and  
 CC screening of therapeutically useful reagents, in gene therapy, for  
 CC chromosome identification, as chromosome marker, and for generating  
 CC probes. An anti-(I)-antibody is useful in diagnostic assays for PRO, e.g.  
 CC detecting its expression in specific cells, tissues or serum, and for  
 CC affinity purification of PRO from recombinant cell culture or natural  
 CC sources. (I) and (II) are useful for tissue typing. This is the amino  
 CC acid sequence of a novel human secreted and transmembrane PRO  
 CC polypeptide.

XX SQ Sequence 182 AA;

Query Match 99.7%; Score 959; DB 6; Length 182;  
 Best Local Similarity 99.5%; Pred. No. 7.6e-95;  
 Matches 181; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPQLGPEAAALRPGWLALLWVSLSCSFSLPASSLSLVFQVRTSYNFGRTFLGLDKC 60  
 DB 1 MEPQLGPEAAALRPGWLALLWVSLSCSFSLPASSLSLVFQVRTSYNFGRTFLGLDKC 60  
 QY 61 NACIGTSICKKPFKEIRSDNWLASHGLPPDSLSYPANYSDDSKINRPVIEFLVSKY 120  
 DB 61 NACIGTSICKKPFKEIRSDNWLASHGLPPDSLSYPANYSDDSKINRPVIEFLVSKY 120  
 QY 121 ONEISDRKICASAPKTCISIRVLKTERFQWQLAKRLTPDLVQDCHQGORELKLFLCM 180  
 DB 121 ONEISDRKICASAPKTCISIRVLKTERFQWQLAKRLTPDLVQDCHQGORELKLFLCM 180  
 QY 181 LR 182  
 DB 181 LR 182

RESULT 15

ADA76122

ID ADA76122 standard; protein; 182 AA.

AC ADA76122;

DT 20-NOV-2003 (first entry)

DE Human PRO polypeptide #86.

XX Human; PRO; secreted polypeptide; transmembrane polypeptide;

KW tumour necrosis factor-alpha; TNF-alpha; chondrocyte cell; tumour;  
 KW cancer; adrenal; lung; colon; breast; prostate; rectum; kidney; cervix;  
 KW liver; microvascular endothelial cell; glucose; FFA;  
 KW skeletal muscle cell; adipocyte cell; pericyte cell;  
 KW inner ear utricular supporting cell; T-lymphocyte cell;  
 KW endothelial cell tube formation; bone disorder; cartilage disorder;  
 KW sports injury; proteoglycan; articular cartilage defect; osteoarthritis;  
 KW rheumatoid arthritis; haemoglobin-associated disorder thalassaemia;  
 KW immune system cell infiltration.

OS Homo sapiens.

XX US2003073212-A.

FN 17-APR-2003.

PD 16-APR-2002; 2002US-00123903.

XX

XX 31-MAR-1997; 97WO-US005230.  
 PR 12-JUN-1998; 98WO-US012456.  
 PR 14-JUL-1998; 98WO-US014552.  
 PR 28-AUG-1998; 98WO-US017888.  
 PR 10-SEP-1998; 98WO-US018824.  
 PR 14-SEP-1998; 98WO-US019093.  
 PR 14-SEP-1998; 98WO-US019094.  
 PR 14-SEP-1998; 98WO-US019177.  
 PR 16-SEP-1998; 98WO-US019330.  
 PR 17-SEP-1998; 98WO-US019437.  
 PR 07-OCT-1998; 98WO-US021141.  
 PR 29-OCT-1998; 98WO-US022991.  
 PR 29-OCT-1998; 98WO-US022992.  
 PR 20-NOV-1998; 98WO-US024855.  
 PR 01-DEC-1998; 98WO-US025108.  
 PR 05-JAN-1999; 99WO-US000106.  
 PR 08-MAR-1999; 99WO-US005028.  
 PR 10-MAR-1999; 99WO-US005190.  
 PR 20-APR-1999; 99WO-US008615.  
 PR 14-MAY-1999; 99WO-US010733.  
 PR 02-JUN-1999; 99WO-US012252.  
 PR 01-SEP-1999; 99WO-US020111.  
 PR 08-SEP-1999; 99WO-US020594.  
 PR 13-SEP-1999; 99WO-US020944.  
 PR 15-SEP-1999; 99WO-US021090.  
 PR 05-OCT-1999; 99WO-US021547.  
 PR 29-NOV-1999; 99WO-US023089.  
 PR 30-NOV-1999; 99WO-US028214.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 01-DEC-1999; 99WO-US028409.  
 PR 01-DEC-1999; 99WO-US028301.  
 PR 01-DEC-1999; 99WO-US028634.  
 PR 02-DEC-1999; 99WO-US028551.  
 PR 02-DEC-1999; 99WO-US028564.  
 PR 16-DEC-1999; 99WO-US028565.  
 PR 20-DEC-1999; 99WO-US030911.  
 PR 20-DEC-1999; 99WO-US030999.  
 PR 22-DEC-1999; 99WO-US030720.  
 PR 30-DEC-1999; 99WO-US031274.  
 PR 30-DEC-1999; 99WO-US031243.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000277.  
 PR 11-FEB-2000; 2000WO-US003565.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 18-FEB-2000; 2000WO-US004342.  
 PR 24-FEB-2000; 2000WO-US004414.  
 PR 24-FEB-2000; 2000WO-US004914.  
 PR 01-MAR-2000; 2000WO-US005004.  
 PR 02-MAR-2000; 2000WO-US005746.  
 PR 10-MAR-2000; 2000WO-US006319.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 20-MAR-2000; 2000WO-US007377.  
 PR 21-MAR-2000; 2000WO-US007532.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 28-JUL-2000; 2000WO-US020710.  
 PR 11-AUG-2000; 2000WO-US022031.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 10-NOV-2000; 2000WO-US030873.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 20-DEC-2000; 2000US-00747259.  
 PR 20-DEC-2000; 2000WO-US034956.  
 PR 28-FEB-2001; 2001US-00796498.



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	86.5	9.0	438	2	T32114	hypothetical prote	
2	83	8.6	398	2	F91202	EpG protein limpo	
3	83	8.6	398	2	A86049	hypothetical prote	
4	81	8.4	424	2	AB1034	UV protection prot	
5	80	8.3	390	2	S85660	hypothetical 43.6K	
6	80	8.3	390	2	B91291	hypothetical prote	
7	80	8.3	390	2	D86132	hypothetical prote	
8	79	8.2	424	2	B38176	samB protein - Sal	
9	79	8.2	985	2	T41135	hypothetical prote	
10	77	8.0	678	2	T50256	probable vacuolar	
11	77	8.0	2374	2	T21052	hypothetical prote	
12	76	7.9	820	2	T26272	hypothetical prote	
13	76	7.9	1006	2	E96683	hypothetical prote	
14	76	7.9	2712	2	T30949	hypothetical prote	
15	76	7.9	3864	2	D87757	protein C4E4.1a [	
16	75	7.8	252	2	G70380	hypothetical prote	
17	75	7.8	386	2	G88765	protein rllB7.1 li	
18	75	7.8	427	2	C86469	protein Fl2K21.14	
19	75	7.8	445	2	T24829	hypothetical prote	
20	74.5	7.7	614	2	T39688	fimbria - fission	
21	74.5	7.7	795	2	T05860	alpha, alpha-trehal	
22	74.5	7.7	1963	2	B98002	IGA-specific metal	
23	74.5	7.7	2004	2	F95133	immunoglobulin A1	
24	74	7.7	398	2	T26274	hypothetical prote	
25	74	7.7	464	2	F37714	hypothetical prote	
26	74	7.7	481	2	T00863	hypothetical prote	
27	74	7.7	600	2	T21231	hypothetical prote	
28	74	7.7	739	2	T25030	hypothetical prote	
29	74	7.7	1109	2	T148536	receptor-like prot	

A; status: preliminary  
A-Molecule type: DNE

100

```

Db      188 ANFYHLGQNPALSGDILKVVYGATFRFDKEALINELDMATVIRQWEEGQRLDPRP 247
      164 --LVQDCHQG 171
      248 RILITGCPIG 257

RESULT 6
B91291
hypothetical protein ECs5298 [imported] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: B91291
R:Havashi, T.; Makino, K.; Ohtsishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A93629; MUID:21156231; PMID:11258796
A:Accession: B91291
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-390 <RAY>
A:Cross-references: GB:BA000007; PIDN:BA038721.1; PID:g13364776; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs5298
C:Superfamily: 2-hydroxyglutaryl-CoA dehydratase beta chain

Query Match      8.3%; Score 80; DB 2; Length 390;
Best Local Similarity 22.1%; Pred. No. 8.6;
Matches 42; Conservative 31; Mismatches 45; Indels 72; Gaps 13;

QY      39 SLVPQVTSYNGRTFLGLDKC-----NACIGTSIC---KKPFKEIRSDNWLASHLGL 89
      83 NLCPLIKSSYGFQKT-----DKCFYFVFDLWVGETTCDGKKQWYE-----YMAE---F 128
      90 PPSLLSYPNYSDDSK--IWRPVEIFRL-----VSKYQNEIS-----DRKIC 130
      129 KPVHVMQLPNSVKDDASRALWK-AEMRLQKAVEERFGHEISEDALDAIALKNRERRAL 187
      131 AS-----ASAPKTCSTIERVLR-----KTERF-----QKWLQAKRLTPD- 163
      188 ANFYHLGQNPALSGDILKVVYGATFRFDKEALINELDMATVIRQWEEGQRLDPRP 247

QY      164 --LVQDCHQG 171
      248 RILITGCPIG 257

RESULT 7
D86132
hypothetical protein yjiM [imported] - Escherichia coli (strain O157:H7, substrain EDL93)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: D86132
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: AB5480; MUID:21074935; PMID:11206551
A:Accession: D86132
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-390 <STO>
A:Cross-references: GB:AE005174; NID:g12519354; PIDN:AA059520.1; GSPDB:GN00145; UWGP:Z59
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yjiM
C:Superfamily: 2-hydroxyglutaryl-CoA dehydratase beta chain

Query Match      8.3%; Score 80; DB 2; Length 390;
Best Local Similarity 22.1%; Pred. No. 8.6;

```

```

Matches 42; Conservative 31; Mismatches 45; Indels 72; Gaps 13;

QY      39 SLVPQVTSYNGRTFLGLDKC-----NACIGTSIC---KKPFKEIRSDNWLASHLGL 89
      83 NLCPLIKSSYGFQKT-----DKCFYFVFDLWVGETTCDGKKQWYE-----YMAE---F 128
      90 PPSLLSYPNYSDDSK--IWRPVEIFRL-----VSKYQNEIS-----DRKIC 130
      129 KPVHVMQLPNSVKDDASRALWK-AEMRLQKAVEERFGHEISEDALDAIALKNRERRAL 187
      131 AS-----ASAPKTCSTIERVLR-----KTERF-----QKWLQAKRLTPD- 163
      188 ANFYHLGQNPALSGDILKVVYGATFRFDKEALINELDMATVIRQWEEGQRLDPRP 247

QY      164 --LVQDCHQG 171
      248 RILITGCPIG 257

RESULT 8
B38176
samA protein - Salmonella typhimurium
C:Species: Salmonella typhimurium
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 20-Jun-2000
C:Accession: B38176
R:Nohmi, T.; Nakura, A.; Nakai, Y.; Watanabe, M.; Murayama, S.Y.; Sofuni, T.
J. Bacteriol. 173, 1051-1063, 1991
A:Title: Salmonella typhimurium has two homologous but different umuDC operons: cloni
A:Reference number: A38176; MUID:91123176; PMID:1991707
A:Accession: B38176
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <NOH>
A:Cross-references: GB:D90202; NID:g217087; PIDN:BA014226.1; PID:g217089
A:Experimental source: strain LT2
C:Genetics:
A:Gene: samB
C:Function:
A:Description: restores UV mutability; involved in mutagenesis
C:Superfamily: umuC protein
C:Keywords: DNA repair; induced mutagenesis; SOS mutagenesis

Query Match      8.2%; Score 79; DB 2; Length 424;
Best Local Similarity 24.2%; Pred. No. 12;
Matches 40; Conservative 22; Mismatches 63; Indels 40; Gaps 6;

QY      23 VSALSCSFLPAS-----SLSLVPOVRTSYNGRTFLGLDKCNACIGTSICKKPFKE 75
      67 VIAFSSNYALYASMSNRVMVHLEELAPRVE-QYSDIMELDIRGIDSCIDFDFGQRLRE 125
      76 ETRSDNWLASHLGLPPDLSLSPANYSDDSKIW-----RPVE 112
      126 HVRSGTGLTIGVGMGTKTAKSAQWA--SKMSQFGGVGLALTLHNQKTEKLLSLQPVE 183
      113 ----IFRLVSKYQNEISDRKICASAPKTC---STERVLRKTER 150
      184 ELWGVGRRISSKILNTMGITLALQLARANTFIRKKNPVVLETRVR 228

QY      76 ETRSDNWLASHLGLPPDLSLSPANYSDDSKIW-----RPVE 112
      126 HVRSGTGLTIGVGMGTKTAKSAQWA--SKMSQFGGVGLALTLHNQKTEKLLSLQPVE 183
      113 ----IFRLVSKYQNEISDRKICASAPKTC---STERVLRKTER 150
      184 ELWGVGRRISSKILNTMGITLALQLARANTFIRKKNPVVLETRVR 228

RESULT 9
T41135
hypothetical protein SPCC1795.08c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T41135
R:Lyne, M.; Rajadream, M.A.; Barrell, B.G.; Oliver, K.; Harris, D.
submitted to the EMBL Data Library, April 1998
A:Reference number: Z21971
A:Accession: T41135
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-985 <LYN>
A:Cross-references: EMBL:AL022598; PIDN:CAA18643.1; GSPDB:GN00068; SPDB:SPCC1795.08c

```

A;Experimental source: strain 972h-; cosmid c1795

C;Genetics:

A;Gene: SPDB:SPAC1795.08c

A;Map position: 3

A;Introns: 3/3; 97/2

Query Match 8.2%; Score 79; DB 2; Length 985;

Best Local Similarity 20.8%; Pred. No. 33;

Matches 41; Conservative 44; Mismatches 74; Indels 38; Gaps 9;

QY 1 MEPLGPEAAALRP--GWL-----ALLWVSALSCSFLPASSLS--LVP--QVRTSY 48

Db 695 VRPPYESSKDIPPEAPLWDEDELLULLLRYSFNWFWASRLTPGLYIPLAKRTAW 754

QY 49 NFGRTFLGDKNCACIGTS-----ICKFKFEIRSDNWLASHLGL-----PPDSLLSYPA 99

Db 755 DCFERWIQVDPRAANVQLTGSARLAQQKLDLSLRSDKVSQHLSDREGTPNHLKHNS 814

QY 100 NY--SDSKTRWPVEIFRLVSKYQNE-----ISDRKICASASA-----PKTCSIE 142

Db 815 YFLFTVSRHYRPITIFEARILKLGREFAKKPTMTKRAIAPSAASTEKUPPVPSPIELS 874

QY 143 RVLKRTERFQKWLQAKR 159

Db 875 RLKSEREAQIQIQIAQR 891

RESULT 10

T50256

probable vacuolar sorting protein [imported] - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 09-Jun-2000 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000

C;Accession: T50256; T39237

R;Skellton, J.; Churcher, C.; Lyne, M.H.; Rajandream, M.A.; Barrell, B.G.; Quail, M.

submitted to the EMBL Data Library, December 1999

A;Reference number: Z25051

A;Accession: T50256

A;Molecule type: DNA

A;Residues: 1-528 <SKE>

A;Cross-references: EMBL:AL133442; PIDN:CAB62830.1; GSPDB:GN00066; SPDB:SPAC767.01c

A;Experimental source: strain 972h(-)

R;Churcher, C.M.; Gentles, S.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1997

A;Reference number: Z21837

A;Accession: T39237

A;Molecule type: DNA

A;Residues: 352-678 <CHU>

A;Cross-references: EMBL:Z98763; PIDN:CAB11498.1; GSPDB:GN00066; SPDB:SPAC9G1.14c

A;Experimental source: strain 972h-; cosmid c9G1

C;Genetics:

A;Gene: SPDB:SPAC9G1.14c; SPDB:SPAC767.01c

A;Map position: 1

C;Superfamily: dynamin-related protein VPS1

Query Match

Best Local Similarity 8.0%; Score 77; DB 2; Length 678;

Matches 34; Conservative 20; Mismatches 53; Indels 38; Gaps 6;

QY 41 VPQVTSYNFRTFLGDKNCACIGTSICKFKFEIRSDNWLASHLGLPPDSLLSYPA 100

Db 542 VPPVETSSSGNQFFG-----SFFGSKNKGSLAAME-----PPPPVLRATT 583

QY 1C1 YSDSKLWRPVEIFRLVSKYQNEISDRKICASAPKTCSTERY-----L 145

Db 584 LSDREK--TDTEVTKLLIMSYFNIV---KRTIADMPKPSISLRMIKYSKEHIQHLEQL 638

QY 146 RKTFRFQKWLQAKRLTPDLVODCHQ 170

Db 639 YKSAFDKLLQSESVTVORKECEQ 663

RESULT 11

T21052

hypothetical protein F22G12.5 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999

C;Accession: T21052; T21281

R;Kershaw, J.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z19365

A;Accession: T21052

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-2374 <WIL>

A;Cross-references: EMBL:Z81066; PIDN:CAB02974.1; GSPDB:GN00019; CESP:F22G12.5

A;Experimental source: clone P17B5

R;Lennard, N.

submitted to the EMBL Data Library, March 1997

A;Reference number: Z19400

A;Accession: T21281

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-2374 <W12>

A;Cross-references: EMBL:Z92831; PIDN:CAB07369.1; GSPDB:GN00019; CESP:F22G12.5

A;Experimental source: clone F22G12

C;Genetics:

A;Gene: CESP:F22G12.5

A;Map position: 1

A;Introns: 47/2; 75/3; 114/3; 172/2; 230/3; 322/1; 376/2; 439/1; 653/2; 731/3; 758/1;

72; 2139/1; 2258/3; 2298/2; 2344/3

Query Match

Best Local Similarity 8.0%; Score 77; DB 2; Length 2374;

Matches 28; Conservative 14; Mismatches 43; Indels 16; Gaps 5;

QY 87 LGLPPDSLLSYPNYSDSKTRWPVEIFRLVSK-----YQNEISDRKICASAPKTC 140

Db 620 LQLEPTTV-----NHRD---LWIVCRVDRILSADTAALYMKSSDPKAVAKLQKTWSM 671

QY 141 IERVLRKTRERFQKWLQAKRLTPDLVQDCHQGRLEKFLCML 181

Db 672 LTRLAHRQRP-AW-TAKLFFELRSNTHQSRARRLECF 710

RESULT 12

T26272

hypothetical protein W07G4.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C;Accession: T26272

R;Baynes, C.

submitted to the EMBL Data Library, August 1996

A;Reference number: Z20187

A;Accession: T26272

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-820 <WIL>

A;Cross-references: EMBL:Z78018; PIDN:CAB01444.1; GSPDB:GN00023; CESP:W07G4.3

A;Experimental source: clone W07G4

C;Genetics:

A;Gene: CESP:W07G4.3

A;Map position: 5

A;Introns: 40/3; 81/3; 122/3; 217/1; 353/2; 577/1; 639/3; 728/1; 778/1

Query Match

Best Local Similarity 7.9%; Score 76; DB 2; Length 820;

Matches 40; Conservative 22; Mismatches 62; Indels 12; Gaps 7;

QY 20 LLWVSA-LSCSFLSPASSLSLVFQVTSYNFGRFTFLGDKNCACIGTSICKFKFEIR 78

Db 696 LLGISAPPAAGSVKPSNLALMNNSSSTNTG---WGFDDPLPTNNSSANKGWDGDM 752

QY 79 SDNWLASHLGLPPDSLLSYPNYSDSKTRWPVEIFRLVSKYQNEISDRKICASAPKT 138

Db 753 GDNMSSNV-LQPTAVLIQ-PTMKDD-KDARRAEQARNEARKEQMDRK--KSAGAKML 807

QY 139 CSIERVLRKTERFQXW 154  
Db 808 ASVE---KKWDDFADW 820

## RESULT 13

E96683  
hypothetical protein F12P19.9 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: E96683  
R:theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizart, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: E96683  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1006 <STO>  
A:Cross-references: GB:AF035173; NID:96227017; PIDN:AAF06053.1; GSPDB:GN00141  
C:Genetics: F12P19.9  
A:Gene: F12P19.9  
A:Map position: 1

Query Match 7.9%; Score 76; DB 2; Length 1006;  
Best Local Similarity 26.7%; Pred. No. 64;  
Matches 31; Conservative 17; Mismatches 48; Indels 20; Gaps 5;

QY 24 SALSCSFSLPASS--LSLIVQVTSYNGRFTGLGKNACIGTSICKKFKKEIRSDN 81  
Db 190 NGLSCSDSFAESDALGPVSSYYETDYFRNS-----DCDRSTGSELCR-----FSSQR 238  
QY 82 WLASHLGLPPDLSLSP---ANYSDDSKIRPVEIFRLVSKYQNEISDRKICASAS 134  
Db 239 FAAS-----PPUSIIITQVTRNSVLKDIMWGAIITGLDGSKNQNDALSPKLESAT 290

## RESULT 14

T30949  
hypothetical protein C44E4.1a - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 22-Oct-1999  
C:Accession: T30949  
R:Sammons, L.; Wohldmann, P.; Gillam, B.  
submitted to the EMBL Data Library, August 1999  
A:Description: The sequence of C. elegans cosmid C44E4.  
A:Reference number: Z20945  
A:Accession: T30949  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2712 <SAM>  
A:Cross-references: EMBL:AF003140; PIDN:AAD47119.1  
A:Experimental source: strain Bristol N2; clone C44E4  
C:Genetics:  
A:Map position: 1  
A:Introns: 44/1; 346/2; 870/1; 1095/3; 1404/3; 1506/2; 1665/1; 1891/3; 1936/3; 2068/2; 2  
A>Note: C44E4.1a

Query Match 7.9%; Score 76; DB 2; Length 2712;  
Best Local Similarity 25.0%; Pred. No. 2.1e+02;  
Matches 27; Conservative 19; Mismatches 34; Indels 28; Gaps 5;

QY 83 LASHLGLPPD-SLISYPANYSDDSKIRPVEIFRLVSKYQNEISDR-----KIC 130  
Db 856 LVSSANPNPTSIMNWEST--SEDSYIIACTDLILLIPQHLQELDRKSVPRDDQWICKLC 914

QY 131 ASASAPKTCSEIRVLRKTERFQXWLAQAKRLTPDLVQDCHQOQRELKFL 178  
Db 915 QLASLSGCSAYR-----QCKKL---LLAMCHNDENKYKIM 947

## RESULT 15

D87757  
protein C44E4.1a [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: D87757  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio.  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_e  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;  
A:Accession: D87757  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-3864 <STO>  
A:Cross-references: GB:chr\_I; PIDN:AAB54167.1; PID:92088725; GSPDB:GN00019; CESP:C44E  
C:Genetics:  
A:Gene: C44E4.1a  
A:Map position: 1

Query Match 7.9%; Score 76; DB 2; Length 3864;  
Best Local Similarity 25.0%; Pred. No. 3.3e+02;  
Matches 27; Conservative 19; Mismatches 34; Indels 28; Gaps 5;

QY 83 LASHLGLPPD-SLISYPANYSDDSKIRPVEIFRLVSKYQNEISDR-----KIC 130  
Db 1993 LVSSANPNPTSIMNWEST--SEDSYIIACTDLILLIPQHLQELDRKSVPRDDQWICKLC 2051  
QY 131 ASASAPKTCSEIRVLRKTERFQXWLAQAKRLTPDLVQDCHQOQRELKFL 178  
Db 2052 QLASLSGCSAYR-----QCKKL---LLAMCHNDENKYKIM 2084

Search completed: June 8, 2004, 16:26:20  
Job time : 21 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: June 8, 2004, 16:20:44 ; Search time 18 Seconds  
(without alignments)  
526.487 Million cell updates/sec

Title: US-10-054-988-114  
Perfect score: 962  
Sequence: 1 MEPQIGPBAALRPQWALL.....DLVQCHQOQRELKFLCMLR 182

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87.5	9.1	316	1	M1AA_PASMU
2	80	8.3	383	1	YJTM_ECOLI
3	79	8.2	424	1	SAMB_SALTY
4	77	8.0	678	1	VPSI_SCHPO
5	76	7.9	366	1	GHSR_HUMAN
6	75.5	7.8	342	1	HXCA_HUMAN
7	75	7.8	152	1	RK34_SPIOL
8	75	7.8	366	1	GHSR_PIG
9	74.5	7.7	422	1	Y140_HUMAN
10	74	7.7	267	1	OVOL_MOUSE
11	73.5	7.6	342	1	HXCA_MOUSE
12	73.5	7.6	1447	1	SGS1_YEAST
13	73.5	7.6	1839	1	POLR_EPMV
14	73	7.6	890	1	K20A_HUMAN
15	73	7.6	1465	1	MYM2_HUMAN
16	72.5	7.5	167	1	UBCE_ARATH
17	72.5	7.5	855	1	ST14_HUMAN
18	72	7.5	424	1	IMPB_SALTY
19	72	7.5	461	1	KDPA_RICPR
20	72	7.5	1020	1	YB03_CAEEL
21	72	7.5	1071	1	SEFI_KLULA
22	72	7.5	1504	1	DPO2_YEAST
23	72	7.5	2206	1	POLG_POL32
24	71.5	7.4	180	1	VG20_BPMV4
25	71.5	7.4	650	1	HS70_CHLRE
26	71.5	7.4	699	1	UVRE_HUMAN
27	71.5	7.4	1280	1	MDRI_HUMAN
28	71	7.4	223	1	BTG4_HUMAN
29	71	7.4	402	1	RENI_RAT
30	71	7.4	424	1	GRF1_HUMAN
31	71	7.4	510	1	YEER_ECOLI
32	70.5	7.3	365	1	VE2_HPV16
33	70.5	7.3	373	1	Y007_METJA

34	70.5	7.3	513	1	CDP2_MAI2B
35	70.5	7.3	642	1	P1MB_YEAST
36	70.5	7.3	1131	1	RFC1_MOUSE
37	70.5	7.3	1212	1	UTY_MOUSE
38	70.5	7.3	1276	1	MDR2_CRIGR
39	70	7.3	371	1	LR3B_ASPNG
40	70	7.3	473	1	YS81_CAEEL
41	70	7.3	533	1	CDP2_ORISA
42	70	7.3	1835	1	DPOL_RHOM6
43	69.5	7.2	505	1	GURA_PYRAE
44	69.5	7.2	648	1	Z202_HUMAN
45	69.5	7.2	1986	1	WA_EMENI

ALIGNMENTS

RESULT 1  
M1AA\_PASMU STANDARD; PRT; 316 AA.  
AC Q9CMC7;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE trna delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPP transferase) (isopentenyl-diphosphate:trna isopentenyltransferase)  
DE (IPTase) (IPPT).  
DE M1AA OR TRPX OR PM0905.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Pw70;  
RX MEDLINE=21145866; PubMed=11248100;  
RA May B J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida Pw70.\*";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).  
CC -!- FUNCTION: Catalyzes the first step in the biosynthesis of 2-methylthio-N6-(delta(2)-isopentenyl)-adenosine (MS[2]i[A]) adjacent to the anticodon of several tRNA species (By similarity).  
CC -!- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA = diphosphate + tRNA containing 6-isopentenyladenosine.  
CC -!- SIMILARITY: Belongs to the IPP transferase family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: A5006129; AAK02989.1; ALT\_INIT.  
CC HMAP; MF 00185; 1.  
CC InterPro: IPR002627; IPPT.  
CC Pfam: PF01715; IPPT; 1.  
CC ProDom: PD004674; IPPT; 1.  
CC TIGRFAMs: TIGR00174; miaA; 1.  
CC Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding; Complete proteome.  
CC NP\_BIND 13  
CC SEQUENCE 316 AA; 35674 MW; C15963CC19269014 CRC64;

Query Match 9.1%; Score 87.5; DB 1; Length 316;  
Best Local Similarity 32.0%; Pred. No. 0.43;  
Matches 32; Conservative 15; Mismatches 32; Indels 21; Gaps 5;  
QY 75 EIRSD-----NWLAHLGLPP-DSLLSYPNYSDSKIMRPVEIFPLVSKYQNEIS 125  
Db 123 EKVRSEIEEKAAQIQWAAALHQLAKIDPLAAQRINPNDQRINRALEFVLTGKSLSELS 182

QY 126 DRK-----ICASASAPKTCISI--ERVLRKTERPQKWLQ 156  
 Db 183 QOKGDSLPYQLQFALADKORSILHDI---ALRFQKMIIE 219

## RESULT 2

YJIM ECOLI  
 ID YJIM ECOLI STANDARD; PRT; 383 AA.  
 AC P39384;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein yjim.  
 GN YJIM OR B4335.  
 OS Escherichia coli.  
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC Enterobacteriaceae; Escherichia.  
 CC NCBI\_TaxID=562;

[1]  
 SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;  
 RA MEDLINE=95334362; PubMed=7610040;  
 RX Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,  
 RA Blattner F.R.;  
 RT "Analysis of the Escherichia coli genome VI: DNA sequence of the  
 region from 92.8 through 100 minutes.";  
 RL Nucleic Acids Res. 23:2105-2119(1995).  
 CC -!- SIMILARITY: STRONG, TO M.JANNA SCHI MJ0007.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U14003; AAA97231.1; ALT\_INIT.

EMBL; AEC00504; AAC77291.1; ALT\_INIT.

DR EcoGene; EGI2574; YJIM.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 383 AA; 42742 MW; 701CBE69D0BFDACC CRC64;

Query Match 8.3%; Score 80; DB 1; Length 383;  
 Best Local Similarity 22.1%; Pred. No. 2.9;  
 Matches 42; Conservative 31; Mismatches 45; Indels 72; Gaps 13;

QY 39 SLVQVQRTSYNFGRTFLGDKC-----NACIGTISIC---KKFFKEIRSDNWLASHLGL 89

Db 76 NLCPLIKSYGFRT-----DKCPYFYSVLVSGETCDGKKWYE-----YMAE---F 121

QY 90 PPDLSLSYPANYSDDSK--IWRPVEIFRL--VSKYQNEIS-----DRKIC 130

Db 122 KPVHWQLFNSVKDDASRALWK-AEMRLQKTVKVEERFGHEISDALRDAIALKNRRRAL 180

QY 131 AS-----ASAPKTCSTVLR-----KTERF-----QKWLQAKRLTPD- 163

Db 181 ANFYHLGQLNPPALSGSDILKVVYGFATFRPDKALINELDMATRVQQWEEGQRLLDRP 240

QY 164 --LVQDCHOG 171

Db 241 RLITGCPIG 250

## RESULT 3

SAMB SALTY

ID SAMB SALTY STANDARD; PRT; 424 AA.

AC F23832;

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE SamB protein.

GN SAMB.

OS Salmonella typhimurium.  
 OG Plasmid 60-mda cryptic.  
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC Enterobacteriaceae; Salmonella.  
 CC NCBI\_TaxID=602;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=LT2;

RX MEDLINE=91123176; PubMed=1991707;

RA Nohmi T., Hakura A., Nakai Y., Watanabe M., Murayama S.Y.,

RA Sofuni T.;

RT "Salmonella typhimurium has two homologous but different umuDC  
 operons: cloning of a new umuDC-like operon (samAB) present in a  
 60-megadalton cryptic plasmid of S. typhimurium.";

RT J. Bacteriol. 173:1051-1063(1991).

RL -!- FUNCTION: Involved in UV protection and mutation.

CC -!- SIMILARITY: Belongs to the DNA polymerase type-Y family.

CC -!- SIMILARITY: Contains 1 umuC domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; D90202; BAA14226.1; -

DR PIR; B38176; B38176.

DR HAMAP; MF\_01113; atypical; 1.

DR InterPro; IPR001126; UmuC\_like.

DR Pfam; PF00817; IWS; 1.

DR PROSITE; PS0173; UmuC; 1.

KW Plasmid; SOS mutagenesis; DNA repair.

FT DOMAIN 2 189

SQ SEQUENCE 424 AA; 47727 MW; PF8C47476CC58A2B CRC64;

Query Match 8.2%; Score 79; DB 1; Length 424;

Best Local Similarity 24.2%; Pred. No. 4.2;

Matches 40; Conservative 22; Mismatches 63; Indels 40; Gaps 6;

QY 23 VSALSCSFLPAS-----SLSSLVFPQVTSYNFQRTFLGLDKCNACIGTISICKKFFKE 75

Db 67 VIATFSNYALXASMSNRVMVHLEELAPRVE-QYSIDEMFLDIRGIDSCIDFDFGRLRE 125

QY 76 ETRSDNWLASHLGLPDSLSLSPANYSDSKIW-----RPVE 112

Db 126 HVRSGTGLTIGVGMGETKTLAKSAQWA--SKWSQFGVIALTLHNQKTEKLLSLQPVE 183

QY 113 ----IPRLVSKYQNEISDRKICASAPKTC---SIERVLKTER 150

Db 184 EIWGVRRISKLNITMGITLQLARANPTFKENFNVLETRVR 228

## RESULT 4

VPS1 SCHPO

ID VPS1 SCHPO STANDARD; PRT; 678 AA.

AC Q9URZ5; O14309;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Vacuolar sorting protein 1.

GN VPS1 OR SPAC767, OIC OR SPAC9G1.14C.

OS Schizosaccharomyces pombe (Fission yeast)

CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

CC Schizosaccharomycetales; Schizosaccharomycetaceae;

CC Schizosaccharomycetes.

CC NCBI\_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

GHRSR HUMAN STANDARD; PRT; 366 AA.

AC Q92847; Q92848, Q96R07;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing  
peptide receptor) (GHRP) (Ghrelin receptor).  
GN GHRS.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).  
RP  
RC TISSUE-Pituitary; PubMed=8688086;  
RX MEDLINE=96337958;  
RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P.,  
RA Liberato P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L.,  
RA Palyha O.C., Anderson J., Paress P.S., Diaz C., Chou M., Liu K.K.,  
RA McKee K.K., Pong S.-S., Chaung L.-Y., Elbrecht A., Dashkevich M.,  
RA Heavens R., Rigby M., Srinathsinghji D.J.S., Dean D.C., Melillo D.G.,  
RA Patchett A.A., Margund R., Griffin P.R., Demartino J.A., Gupta S.K.,  
RA Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;  
RT \*A receptor in pituitary and hypothalamus that functions in growth  
RT hormone release.";  
RL Science 273:974-977 (1996).  
[2]  
RN SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).  
RP  
RX MEDLINE=21255649; PubMed=11356716;  
RA Peterenn S., Rasch A.C., Penshorn M., Beil F.U., Schulte H.M.;  
RT "Genomic structure and transcriptional regulation of the human growth  
RT hormone secretagogue receptor.";  
RL Endocrinology 142:2649-2659 (2001).  
[3]  
RN SEQUENCE FROM N.A. (ISOFORM 1B).  
RA Kopatz S.A., Aronstam R.S., Sharma S.V.;  
RT "cDNA clones of human proteins involved in signal transduction  
RT sequenced by the Guthrie cDNA resource center (www.cdna.org).";  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
[4]  
RN FUNCTION.  
RP  
RX MEDLINE=21219832; PubMed=11322507;  
RA Smith R.G., Leonard R., Bailey A.R.T., Palyha O.C., Feighner S.D.,  
RA Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;  
RT "Growth hormone secretagogue receptor family members and ligands.";  
RL Endocrine 14:9-14 (2001).  
[5]  
RN FUNCTION.  
RP  
RX MEDLINE=20067959; PubMed=10604470;  
RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;  
RT "Ghrelin is a growth-hormone-releasing acylated peptide from  
RT stomach.";  
RL Nature 402:656-660 (1999).  
CC -!- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.  
CC Stimulates growth hormone secretion. Binds also other growth  
CC hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)  
CC as well as non-peptide, low molecular weight secretagogues (e.g.  
CC L-692,429, MK-0677, adenosine).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=2;  
CC Name=1A;  
CC IsoId=Q92847-1; Sequence=Displayed;  
CC Name=1B;  
CC IsoId=Q92847-2; Sequence=VSP\_001916, VSP\_001917;  
CC -!- TISSUE SPECIFICITY: Pituitary and hypothalamus.  
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way



DR PROSITE; PS00027; HOMEOBOX 1; 1.  
 DR PROSITE; PS00071; HOMEOBOX 2; 1.  
 XW HOMEOBOX; DNA-binding; Developmental protein; Nuclear protein;  
 XW Transcription regulation.

FT DNA\_BIND 268 327 HOMEOBOX.  
 FT CONFLICT 118 118 K -> N (IN REF. 1).  
 FT CONFLICT 265 265 A -> G (IN REF. 3).  
 FT CONFLICT 271 271 MISSING (IN REF. 3).  
 SQ SEQUENCE 342 AA; 38072 MW; BD8127FD43C2A37B CRC64;

Query Match 7.8%; Score 75.5; DB 1; Length 342;  
 Best Local Similarity 26.7%; Pred. No. 7;  
 Matches 43; Conservative 15; Mismatches 78; Indels 25; Gaps 5;

QY 6 GPERAALPGWALLLWLSALSCSPSLPASSLSLVQVRYNFGRTPLGLDKNACIG 65  
 DB 123 GPERALY-----SHPLPESCLGEHEVPVPSYRASPSYALDKTPHCSSG 166  
 QY 66 TSICKKFFKEIRSDNWLASHGLPP-DSLLSYPNKSYDDSKIWPRVIFRLVS-----K 113  
 DB 167 ANDPEAPF-EQASLNPRAEHLESQLGKGVSPETPKSDSYTPSPNEIKTEQSLAGPKG 225  
 QY 120 YQNEISDRKICASAPKTCSTIE--RVLRKTERFQKWLQAK 158  
 DB 226 SPSESEKERAKAADSSPOTSUNEAKEIKAEITGNWLTAK 266

## RESULT 7

RK34\_SPTOL STANDARD; PRT; 152 AA.  
 AC P82244;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 50S ribosomal protein L34, chloroplast precursor.  
 GN RPL34.  
 OS Spinacia oleracea (Spinach).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;  
 OC Caryophyllales; Amaranthaceae; Spinacia.  
 OX NCBI\_TaxID=3562;  
 RN [1]  
 RP SEQUENCE FROM N.A., SEQUENCE OF 92-101, AND MASS SPECTROMETRY.  
 RC STRAIN=cv. Alvaro; TISSUE=Leaf;  
 RX MEDLINE=20435798; PubMed=10874046;  
 RA Yamaguchi K., Subramanian A.R.;  
 RT "The plastid ribosomal proteins. Identification of all the proteins in  
 the 50 S subunit of an organelle ribosome (chloroplast).";  
 RL J. Biol. Chem. 275:28466-28482(2000).  
 CC -!- FUNCTION: This protein binds directly to 23S ribosomal RNA.  
 CC -!- SUBCELLULAR LOCATION: Chloroplast.  
 CC -!- TISSUE SPECIFICITY: Ubiquitous.  
 CC -!- MASS SPECTROMETRY: YW-6787.1; METHOD=Electrospray; RANGE=92-152.  
 CC -!- SIMILARITY: Belongs to the L34 family of ribosomal proteins.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF238221; AAF64157.1; -;  
 DR InterPro; IPR000271; Ribosomal L34.  
 DR Pfam; PF00468; Ribosomal L34; 1.  
 DR TIGRFAMs; TIGR01030; rpmh bact; 1.  
 DR PROSITE; PS00784; RIBOSOMAL L34; FALSE NEG.  
 KW Ribosomal protein; Chloroplast; Transl. peptide; rRNA-binding.  
 FT TRANSIT 1 91 CHLOROPLAST.  
 FT CHAIN 92 152 50S RIBOSOMAL PROTEIN L34.  
 SQ SEQUENCE 152 AA; 16095 MW; 755A99D0441AD818 CRC64;

Query Match 7.8%; Score 75; DB 1; Length 152;  
 Best Local Similarity 40.8%; Pred. No. 2.9;  
 Matches 20; Conservative 5; Mismatches 14; Indels 10; Gaps 2;

QY 25 ALSCSFSLPASSLSLVQVRYNFGRTPLGLD-KONACIGTSICKKF 72  
 DB 48 SLHCSFLAPSSLSL-----NFSGLSLGLDLTSNTGVSTDRCRRF 87

## RESULT 8

GHSR\_PIG STANDARD; PRT; 366 AA.  
 AC Q95254; Q95255;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing  
 peptide receptor) (GHRP) (Ghrelin receptor).  
 GN GHSR.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).  
 RC STRAIN=Yorkshire; TISSUE=Pituitary;  
 RX MEDLINE=96337998; PubMed=868086;  
 RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P.,  
 RA Liberator P.A., Rosenblum J., Pares F.S., Diaz C., Chou M., Liu K.K.,  
 RA Palyha O.C., Anderson J., Pares F.S., Hamelin M., Hreniuk D.L.,  
 RA McKee K.K., Pong S.-S., Chung L.-Y., Elbrecht A., Daskiewicz M.,  
 RA Heavens R., Rigby M., Sirinathsinghji D.J.S., Dean D.C., Melillo D.G.,  
 RA Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K.,  
 RA Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;  
 RT "A receptor in pituitary and hypothalamus that functions in growth  
 hormone release".  
 RL Science 273:974-977(1996).  
 CC -!- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.  
 CC Stimulates growth hormone secretion. Binds also other growth  
 hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)  
 CC as well as non-peptide, low molecular weight secretagogues (e.g.  
 CC L-692,429, MK-0677, adenosine).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1A;  
 CC IsoId=Q95254-1; Sequence=Displayed;  
 CC Name=1B;  
 CC IsoId=Q95254-2; Sequence=VSP\_001918, VSP\_001919;  
 CC -!- TISSUE SPECIFICITY: Pituitary and hypothalamus.  
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U60178; AAC48630.1; -;  
 DR EMBL; U60180; AAC48631.1; -;  
 DR InterPro; IPR000276; GPCR\_Rhodpen.  
 DR Pfam; PF00001; 7tm 1; 1.  
 DR PRINTS; PR00237; GPCRHHODPSN  
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.  
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL 2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Alternative splicing.  
 FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 41 66 1 (POTENTIAL).  
 FT DOMAIN 67 72 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 73 96 2 (POTENTIAL).

FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 118 139 3 (POTENTIAL).  
 FT DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 163 183 4 (POTENTIAL).  
 FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 212 235 5 (POTENTIAL).  
 FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 264 285 6 (POTENTIAL).  
 FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 303 326 7 (POTENTIAL).  
 FT DOMAIN 327 366 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 116 198 BY SIMILARITY.  
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 266 289 AVTVFAPILCWLFHVGRLYFSSK -> GGSQCAULESLPG  
 FT PLHSSCLFSSP (in isoform 1B).  
 FT /FTid=VSP\_001918.  
 FT Missing (in isoform 1B).  
 FT /FTid=VSP\_001919.  
 FT VARSPLIC 290 366  
 FT SEQUENCE 366 AA; 41194 MW; 2C850B3EF61B7C1C CRC64;

Query Match 7.8%; Score 75; DB 1; Length 366;  
 Best Local Similarity 25.6%; Pred. No. 8.5;  
 Matches 39; Conservative 23; Mismatches 46; Indels 48; Gaps 8;  
 QY 11 ALRPGWLALLWVSALSCFSLPSSISLVPQVTSYNGRTFLGDKCNACIGTSICK 70  
 Db 204 AVRSGLLTVMVWS--SVFFLFVCLTVLSLI-----GRKLWRRKGEAAVGSRLD 255  
 QY 71 KPFEKIRSDN-----WLASHLG-----LPDS-----LLSY----- 97  
 Db 256 QNHQKVVMVLAVVVFAPILCWLFHVGRLYFSSLEPGSVFELAQISQYCNLSVFLYLS 315  
 QY 98 -----PANSYSDSKLWVPEIRLVSKYQNEISDRKI 129  
 Db 316 AAINPLLYNIMSKYR-VAVFKLLG--FEPFSQKRL 348

RESULT 9  
 Y140 HUMAN  
 ID Y140 HUMAN STANDARD; PRT; 422 AA.  
 AC Q14153;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein KIAA0140.  
 GN KIAA0140.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow;  
 RX MEDLINE=96127530; PubMed=8530280;  
 RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;  
 RT "Prediction of the coding sequences of unidentified human genes. IV.  
 RT analysis of cDNA clones from human cell line KG-1.";  
 RL DNA Res. 2:167-174 (1995).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D50930; BAA09489.1; --  
 KW Hypothetical protein.  
 SQ SEQUENCE 422 AA; 45796 MW; 8F607F093408CEAC CRC64;

Query Match 7.7%; Score 74.5; DB 1; Length 422;  
 Best Local Similarity 23.3%; Pred. No. 11;  
 Matches 30; Conservative 19; Mismatches 45; Indels 35; Gaps 5;  
 QY 33 PASSLSLVPQVTSYNGRTFLGDKCNACIGTSICKPKPKKEIRSDNWLASHLG---L 89  
 Db 63 PSTSIWECLPEKDS-----LHREAVTACAVTSLIK-----DLSIDHNGPNSA 107  
 QY 90 PPDLSLSYPANYSD-----SKWRPVEIRFL-----VSKYQNEISDRKICAS 132  
 Db 108 PPSKRCRSLSFSDMSCHTSWRPLGSKVWTPVEKRCYSGGSVQRYSGFSTMQRSS 167  
 QY 133 ASAPKTCST 141  
 Db 168 FSLPSRANV 176

RESULT 10  
 OV01 MOUSE  
 ID OV01 MOUSE STANDARD; PRT; 267 AA.  
 AC Q9WTJ2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Putative transcription factor Ovo-like 1 (mOvol) (mOvola).  
 GN OVOLA.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RA Schonbaum C.P., Fantes J., Mahowald A.P.;  
 RT "Characterization of mouse and Caenorhabditis elegans genes related to  
 RT the Drosophila melanogaster ovo/svb gene";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 FUNCTION.  
 RX MEDLINE=99026118; PubMed=9808631;  
 RA Dai X., Schonbaum C., Degenstein L., Bai W., Mahowald A., Fuchs E.;  
 RT "The ovo gene required for cuticle formation and oogenesis in flies is  
 RT involved in hair formation and spermatogenesis in mice.";  
 RL Genes Dev. 12:3452-3463 (1998).  
 CC -|- FUNCTION: Putative transcription factor. Involved in hair  
 CC formation and spermatogenesis. May function in the differentiation  
 CC and/or maintenance of the urogenital system.  
 CC -|- SUBCELLULAR LOCATION: Nuclear.  
 CC -|- TISSUE SPECIFICITY: EXPRESSED IN SKIN, TESTIS, KIDNEY AND WEAKLY  
 CC IN LUNG. NOT DETECTED IN HEART, BRAIN, SPLEEN, LIVER AND SKELETAL  
 CC MUSCLE.  
 CC -|- DEVELOPMENTAL STAGE: FIRST EXPRESSED AT E14.5 DAY IN THE  
 CC SUPRABASAL LAYERS OF DEVELOPING EPIDERMIS, AT E15.5 EXPRESSION  
 CC BEGINS IN THE INNER CELLS OF DEVELOPING HAIR GERMS AND RESTRICTED  
 CC TO INNER ROOT SHEATH AND/OR PRECORTICAL CELLS OF DEVELOPING HAIR  
 CC FOLLICLES.  
 CC -|- SIMILARITY: Contains 4 C2H2-type zinc fingers.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF134804; AAD29689.1; --  
 DR EMBL; AF134805; AAD29690.1; --  
 DR HSSP; P07248; 2ADR.  
 DR MGD; MGI:1330290; Ovov1.  
 DR GO; GO:0005634; C:nucleus; IDA.  
 DR GO; GO:0008544; P:epidermal differentiation; IMP.  
 DR GO; GO:0007498; P:mesoderm development; IMP.

DR GO:0007283; P:permatogenesis; IMP.  
 DR InterPro: IPR007087; Znf\_C2H2.  
 DR Pfam: PF00096; zf\_C2H2; 4.  
 DR ProDom: PD000003; Znf\_C2H2; 1.  
 DR SMART: SMC0355; Znf\_C2H2; 4.  
 DR PROSITE: PS00028; ZINC FINGER C2H2\_1; 3.  
 DR PROSITE: PS00157; ZINC FINGER C2H2\_2; 3.  
 DR Zinc-finger; Metal-binding; DNA-binding; Nuclear protein; Repeat;  
 KW Transcription regulation.  
 FT ZN\_FING 118 140 C2H2-TYPE 1.  
 FT ZN\_FING 146 168 C2H2-TYPE 2.  
 FT ZN\_FING 174 197 C2H2-TYPE 3.  
 FT ZN\_FING 213 236 C2H2-TYPE 4.  
 SQ SEQUENCE 267 AA; 30221 MW; 30221 MW; DAD4F51150C21C2D CRC64;  
 Query Match 7.7%; Score 74; DB 1; Length 267;  
 Best Local Similarity 21.6%; Pred. No. 7.2;  
 Matches 37; Conservative 23; Mismatches 81; Indels 30; Gaps 4;  
 QY 2 EPQGPAAALRPGLALLLVVSALSCSFS-----LPASSLSLVFPQVTSYNG 51  
 DB 42 QYRPEASVAPPSPLADSLRSSYSVAPGCVVQLPSEDSVSHLTDPQSRDQGL 101  
 QY 52 RTFLGLDKNCACIGTS-----ICKKFFKEEIRSDNWLASHLGLPPDPSLSLYEANY 102  
 DB 102 RT-----KMKVTLGDSPNGDLFTCHI:CKKSFTHQRLNRMKCHNDVKHLCYCGKGF 156  
 QY 103 DSKTWRPVEIFRLSKYQNEISDRKICASAPKTCSTIRVLKTERFQK 153  
 DB 157 TDFDLKRRVTRHTGVRPKYSLCDK-----AFTQRCLESHLKKIHGQQ 201

## RESULT 11

HKCA\_MOUSE  
 ID HKCA\_MOUSE STANDARD; PRT; 342 AA.  
 AC P31257;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Homeobox protein Hox-C10 (Hox-3.6).  
 GN HOXC10 OR HOXC-10 OR HOX-3.6.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CD-1;  
 RX MEDLINE=92360500; PubMed=1353983;  
 RA Peterson R.L.; Jacobs D.F.; Awgulewitsch A.;  
 RT "Hox-3.6: isolation and characterization of a new murine homeobox  
 RL gene located in the 5' region of the Hox-3 cluster.";  
 RN [2]  
 RP Mech. Dev. 37:1151-1166 (1992).  
 RP SEQUENCE OF 289-313 FROM N.A.  
 RX MEDLINE=92073357; PubMed=1720547;  
 RA Murtha M.T.; Leckman J.F.; Ruddle F.H.;  
 RT "Detection of homeobox genes in development and evolution.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10711-10715 (1991).  
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of  
 CC a developmental regulatory system that provides cells with  
 CC specific positional identities on the anterior-posterior axis.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: Belongs to the Abd-B homeobox family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

DR EMBL; X63507; -, NOT ANNOTATED\_CDS.  
 DR EMBL; M81658; AAA63310.1; -.  
 DR PIR; A56552; A56552.  
 DR HGSP; P02833; 9ANT.  
 DR TRANSFAC; T03350; -.  
 DR MGD; MGI:96192; Hoxc10.  
 DR InterPro: IPR001356; Homeobox.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00024; HOMEBOX.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS00071; HOMEBOX\_2; 1.  
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;  
 FT Transcription regulation.  
 FT DNA\_BIND 268 327 HOMEBOX.  
 SQ SEQUENCE 342 AA; 38195 MW; 7BB2E117A768F52B CRC64;  
 Query Match 7.6%; Score 73.5; DB 1; Length 342;  
 Best Local Similarity 25.3%; Pred. No. 11;  
 Matches 42; Conservative 15; Mismatches 74; Indels 35; Gaps 6;  
 QY 6 GPERAALRPGLALLLVVSALSCSFS-----SHPSPSCGHEHVPSPYRASFSALOKTPHCAG 65  
 DB 123 GPERALY-----SHPSPSCGHEHVPSPYRASFSALOKTPHCAG 166  
 QY 66 TSICKKFFKEEIRSDNWLASHLGLPP-DSLISYPANYSDSKIWRPVEIFRLSKYQNEI 124  
 DB 167 ANEFAPPE-EDRASLNPRTEHLESFOLGKGVFPETPKSDSQTPSPNEI-----KTEQSL 220  
 QY 125 SDRKICASA-----SAPKTCISIR--RVLRKTERFQKWLQAK 158  
 DB 221 AGPKASPSESEKERAKTADSPDTSNDAKEIKAEKTKAENTGNWLTAK 266

## RESULT 12

SGS1\_YEAST  
 ID SGS1\_YEAST STANDARD; PRT; 1447 AA.  
 AC P35187;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Helicase SGS1 (Helicase TPS1).  
 GN SGS1 OR TPS1 OR YMR190C OR YMR646.02C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=W303;  
 RX MEDLINE=95059068; PubMed=7969174;  
 RA Gangloff S.; McDonald J.P.; Bendixen C.; Arthur L.; Rothstein R.;  
 RT "The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA  
 RL helicase homolog: a potential eukaryotic reverse gyrase.";  
 RN [2]  
 RP Mol. Cell. Biol. 14:8391-8398 (1994).  
 RP SEQUENCE FROM N.A.  
 RA Romeo A.M.; Kleff S.; Sternglanz R.;  
 RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RX MEDLINE=97313268; PubMed=9169872;  
 RA Bowman S.; Churcher C.M.; Badcock K.; Brown D.; Chillingworth T.;  
 RA Connor R.; Dedman K.; Devlin K.; Gentles S.; Hamlin N.; Hunt S.;  
 RA Jagels K.; Lye G.; Moule S.; Odell C.; Pearson D.; Rajandream M.A.;  
 RA Rice P.; Skelton J.; Walsh S.; Whitehead S.; Barrell B.G.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome  
 RL XIII.";  
 RL Nature 387:90-93 (1997).  
 RN [4]  
 RP STRUCTURE BY NMR OF HRDC DOMAIN.

```

RX MEDLINE=20113483; PubMed=10647186;
RA Liu Z., Macias M.J., Bottomley M.J., Sier G., Lange J.P., Nilges M.,
RA Bork P., Sattler M.;
RT "The three-dimensional structure of the HRDC domain and implications
RT for the Werner and Bloom syndrome proteins.";
RL Structure 7:1557-1566(1999).
CC -!- FUNCTION: Interacts with topoisomerases II and TOP3. Could create
CC a deleterious topological substrate that TOP3 preferentially
CC resolves. The TOP3-SGS1 protein complex may function as a
CC eukaryotic reverse gyrase introducing positive supercoils into
CC extrachromosomal ribosomal DNA rings.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the helicase family. RecQ subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L07870; AAA35167.1; -
CC EMBL: U22341; AAB60289.1; -
CC EMBL: Z47815; CA87811.1; -
CC PIR: S50918; S50918.
CC PDB: 1D8B; 10-JAN-00.
CC GerMOnline: 142864; -
CC SGD: S0004802; SGS1.
CC GO: GO:0005730; C:nucleolus; IDA.
CC GO: GO:0004003; F:ATP dependent DNA helicase activity; IDA.
CC GO: GO:0007001; P:Chromosome organization and biogenesis {sen. .; IMP.
CC GO: GO:0006268; P:DNA unwinding; IDA.
CC GO: GO:0045132; P:meiotic chromosome segregation; IMP.
CC GO: GO:0000070; P:mitotic chromosome segregation; IMP.
CC InterPro: IPR001410; DEAD.
CC InterPro: IPR002464; DEAD box.
CC InterPro: IPR001650; Helicase_C.
CC InterPro: IPR002121; HRDC.
CC InterPro: IPR004589; RecQ.
CC Pfam: PF00270; DEAD; 1.
CC Pfam: PF00271; helicase_C; 1.
CC Pfam: PF00570; HRDC; 1.
CC SMART: SM00487; DEXDC; 1.
CC SMART: SM00490; HELIC_C; 1.
CC SMART: SM00341; HRDC; 1.
CC TIGRfam: TIGR00614; recQ fam; 1.
CC PROSITE: PS00690; DEAD_ATP_HELICASE; UNKNOWN 1.
KW Helicase; ATP-binding; Nuclear protein; 3D-structure.
FT DOMAIN 631 639 ASP/GLU-RICH (HIGHLY ACIDIC).
FT NP BIND 714 721 ATP (BY SIMILARITY).
FT SITE 808 811 DEAD BOX.
SQ SEQUENCE 1447 AA; 163836 MW; 0DC320B41756A3C3 CRC64;

Query Match 7.6%; Score 73.5; DB 1; Length 1447;
Best Local Similarity 22.3%; Pred. No. 64;
Matches 47; Conservative 28; Mismatches 81; Indels 55; Gaps 10;

QY 1 MEPQLGPEAAALRPGWALLLVWSALSCSPSPASSLSLVPOVTSYNF--GRFTPLGLD 58
DB 934 MEPD---ERLSVQKAWQADEIQVICATVAFGMDK-----PDVRFVYHFTVPTLEGY 985

QY 59 KNCACIGT---SICKKPF-----KEIRSDNWL-----ASHGLPDPDLSLYPANY 101
DB 986 QETGRAGRDGNYSCITVPSFDITMTQTIQKXLDRENKEHLN-KLQVMAVCNV 1044

QY 102 SDDSKIMRPVEIFRLVSKYQNEISDRKICASAPKTC-----SIRVLRLK 147
DB 1045 TDCRR-----KLVSFYFEDFDSKLC-----HKNCDCNCSANVINEEDVTEPAK 1091

QY 148 TERFQKWLQAKRLTPOLVQDCHQORELKFL 178
DB 1092 IVKLVESIQNERKVTIIYQDVFEGKRSRKIV 1122

```

```

RESULT 13
POLR EPMV STANDARD; PRT; 1839 AA.
ID P20126;
AC 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RNA replicase polypeptide (EC 2.7.7.48).
OS Eggplant mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymoviridae;
OC Tymovirus.
OC NCBI_TaxID=12151;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=90021185; PubMed=2800336;
RA Osorio-Keese M.E., Keese P., Gibbs A.;
RT "Nucleotide sequence of the genome of eggplant mosaic tymovirus.";
RL Virology 172:547-554(1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J04374; AAA43039.1; -
CC MEROPS: C21.001; -
CC InterPro: IPR008043; Peptidase C21.
CC InterPro: IPR001788; RNA dep RNAPol2.
CC InterPro: IPR007095; RNA_pol_DS_PS.
CC InterPro: IPR007094; RNA_pol_PSVir.
CC InterPro: IPR000606; Viral_helicase1.
CC Pfam: PF05381; Peptidase C21; 1.
CC Pfam: PF00978; RNA dep RNAPol2; 1.
CC Pfam: PF01443; Viral_helicase1; 1.
CC Transferase; RNA-directed RNA polymerase; Polyprotein; ATP-binding.
FT NP BIND 965 972 ATP (BY SIMILARITY).
SQ SEQUENCE 1839 AA; 204731 MW; FD8DC1F115E7861 CRC64;

Query Match 7.6%; Score 73.5; DB 1; Length 1839;
Best Local Similarity 20.1%; Pred. No. 86;
Matches 28; Conservative 23; Mismatches 49; Indels 39; Gaps 4;

QY 79 SDNWLASHL-----GLPPDSLALSY-----ANYSDDS 105
DB 1423 SDQLLGQHLFSSLCIAYGNPNFVFPQPELSEICINDYALQSSKTOATIVANQRSD 1482

QY 106 KIRPVEIFRLVSKYQNEISDRKICASASAPKTCSTER-----VLRKTRFKWLQAKRL 160
DB 1483 PDWRLTAV-RIPAKAQHKVNDASIFSGWKACQTLALMHGVIILVLPVKKYQIFDKDR 1541

QY 161 TPDLVQDCHQORELKFLC 179
DB 1542 PPHIYHCGKTPSOLSONC 1560

RESULT 14
K20A HUMAN
ID K20A HUMAN STANDARD; PRT; 890 AA.
AC O95235;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kinesin family member 20A (Rabkinesin-6) (Rab6-interacting kinesin-
DE like protein) (GG10.2).
GN KIF20A OR RAB6KIFL.
OS Homo sapiens (Human).

```



```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Endothelial cells;
RX MEDLINE=99252096; PubMed=10233894;
RA Horrevorts A.J.G., Fontijn R.D., van Zonneveld A.J., de Vries C.J.M.,
RA ten Cate J.W., Pannkoek H.;
RT "vascular endothelial genes that are responsive to tumor necrosis
RT factor-alpha in vitro are expressed in atherosclerotic lesions,
RT including inhibitor of apoptosis protein-1, stannin, and two novel
RT genes.";
RL Blood 93:3418-3431(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=20267854; PubMed=10806357;
RA Lai F., Fernald A.A., Zhao N., Le Beau M.M.;
RT "cDNA cloning, expression pattern, genomic structure and chromosomal
RT location of RAB6KIFL, a human kinesin-like gene.";
RL Gene 248:117-125(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares V.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.S., Toshlyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bock S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.M.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: INTERACTS WITH GUANOSINE TRIPHOSPHATE (GTP)-BOUND FORMS
CC OF RAB6A AND RAB6B. MAY ACT AS A MOTOR REQUIRED FOR THE RETROGRADE
CC RAB6 REGULATED TRANSPORT OF GOLGI MEMBRANES AND ASSOCIATED
CC VESICLES ALONG MICROTUBULES. HAS A MICROTUBULE PLUS END-DIRECTED
CC MOBILITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Golgi (By similarity).
CC -!- SIMILARITY: Belongs to the kinesin-like protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF070672; AAC82320.1; -
CC EMBL; AF153329; AAD37806.1; -
CC EMBL; BC012999; AAH12999.1; -
CC HSSP; P17119; 3KAR.
CC Genew; HGNC:9787; KIF20A.
CC MIM; 605664; -
CC GO; GO:0005794; C:Golgi apparatus; TAS.
CC GO; GO:0005215; F:transporter activity; TAS.
CC GO; GO:0016192; P:vesicle-mediated transport; TAS.
CC InterPro; IPR001752; kinesin_motor.
CC Pfam; PF00225; kinesin; 1.

```

---

```

DR PRINTS; PRO0380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; FALSE_NEG.
DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Microtubule; ATP-binding; Coiled coil; Golgi stack;
KW Protein transport; Transport.
FT DOMAIN 70 KINESIN-MOTOR.
FT NP_BIND 160 167 ATP (POTENTIAL).
FT DOMAIN 611 762 COILED COIL (POTENTIAL).
FT DOMAIN 763 890 GLOBULAR (POTENTIAL).
SQ SEQUENCE 890 AA; 100277 MW; 5620264615496051 CRC64;
Query Match 7.6%; Score 73; DB 1; Length 890;
Best Local Similarity 20.0%; Pred. No. 40;
Matches 32; Conservative 26; Mismatches 46; Indels 56; Gaps 6;
QY 68 ICKKFKFKIRSDNWLASHGLGLPPDLSLSPYANYSDDSKIRPVEIFRLVSKYQNEISDR 127
DB 592 ICNEMVEQMQOREQWCSEHLDYCKELL--EEMVEKLNILKE---SLTSFYQEEIQER 644
QY 128 -----KICASASAPKTSIERVLAK----- 147
DB 645 DEKIEELALLQEARQCSVAHQSGSELALRRSRQLAASASTQQLQEVKAKLQCKKAEIN 704
QY 148 --TE---RQKWLQ----AKLTPDLVQDCHQORELKEF 178
DB 705 STTEELHKYQKMLPSPPSAKPFTIDVDKKLEGGKXNIRLL 744
RESULT 15
MYM2_HUMAN STANDARD; PRT; 1465 AA.
ID MYM2_HUMAN
AC P54296;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myomesin 2 (M-protein) (165 kDa titin-associated protein) (165 kDa
DE connectin-associated protein).
OS Homo sapiens (Human).
GN Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Skeletal muscle;
RX MEDLINE=94095665; PubMed=7505783;
RA Vinkewiler U., Obermann W., Weber K., Fuerst D.O.;
RT "The globular head domain of titin extends into the center of the
RT sarcomeric M band. cDNA cloning, epitope mapping and immunoelectron
RT microscopy of two titin-associated proteins.";
RL J. Cell Sci. 106:319-330(1993).
CC -!- FUNCTION: Major component of the vertebrate myofibrillar M band.
CC Binds myosin, titin, and light meromyosin. This binding is dose
CC dependent.
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 5 fibronectin type III domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X69089; CAA4832.1; -
CC PIR; S43529; S43529.
CC HSSP; P40189; 1BQU.
CC Genew; HGNC:7614; MYOM2.
CC MIM; 603509; -
CC GO; GO:0008107; F:structural constituent of muscle; TAS.
CC GO; GO:0006936; P:muscle contraction; TAS.

```

```
DR InterPro: IPR008957; FN_III-like.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003962; FNIII_subd.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig_c2.
DR Pfam: PF00041; fn3; 5.
DR Pfam: PF00047; ig; 4.
DR PRINTS: PR00014; ENTYPETIII.
DR SMART: SM00060; FN3; 3.
DR SMART: SM00408; IGC2; 1.
DR PROSITE: PS00835; IG_LIKE; 5.
KW Immunoglobulin domain; Muscle protein; Thick filament; Repeat.
FT DOMAIN 154 245 IG-LIKE C2-TYPE 1.
FT DOMAIN 266 371 IG-LIKE C2-TYPE 2.
FT DOMAIN 383 469 FIBRONECTIN TYPE-III 1.
FT DOMAIN 511 597 FIBRONECTIN TYPE-III 2.
FT DOMAIN 612 696 FIBRONECTIN TYPE-III 3.
FT DOMAIN 711 797 FIBRONECTIN TYPE-III 4.
FT DOMAIN 813 899 FIBRONECTIN TYPE-III 5.
FT DOMAIN 904 1002 IG-LIKE C2-TYPE 3.
FT DOMAIN 1130 1211 IG-LIKE C2-TYPE 4.
FT DOMAIN 1345 1434 IG-LIKE C2-TYPE 5.
SQ SEQUENCE 1465 AA; 164793 MW; 6F2927BSA1E69F2D CRC64;

Query Match 7.6%; Score 73; DB 1; Length 1465;
Best Local Similarity 29.0%; Pred. No. 73;
Matches 38; Conservative 15; Mismatches 50; Indels 28; Gaps 9;

QY 51 GRTFLGLDKNACIGTS---ICKKPKKEIRSD-----NWLASHLGL---PPD--SLSY 97
Db 924 GNIYLGFD-CQEMTDASQFTWCKSY--EEISDDERFKIETVGDHSLKYLKNDKRDIGTY 983
QY 98 PANYSD-----DSKIWRPVIEIPELVSKYONEISDRKICASASAPKTCSEIIVLTKTERFQ 152
Db 981 SVSVSDTDGVSSSFVLDPELERLMA-LSNEIKNPTI-----PLKSELAYEIFDKGRVR 1033
QY 153 KWLQAKRLTPD 163
Db 1034 FWLQAEHLSPD 1044
```

Search completed: June 8, 2004, 16:24:53  
Job time : 19 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2004, 16:21:44 ; Search time 41 Seconds  
(without alignments)  
1400.593 Million cell updates/sec

Title: US-10-054-988-114  
Perfect score: 962  
Sequence: 1 MEPQLGPEAAALRPGMLALL.....DLVQDCHQGQRELKFLQMLR 182

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_mhc:\*
- 9: sp\_organelle:\*
- 10: sp\_phase:\*
- 11: sp\_plant:\*
- 12: sp\_rodent:\*
- 13: sp\_virus:\*
- 14: sp\_verticbrate:\*
- 15: sp\_unclassified:\*
- 16: sp\_virus:\*
- 17: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	962	100.0	182	4 Q9H7Y0	Q9H7Y0 homo sapien
2	868	90.2	166	4 Q8WX00	Q8WX00 homo sapien
3	655	68.1	435	11 Q8C3I9	Q8C3I9 mus musculus
4	120	12.5	430	4 Q8NDZ4	Q8NDZ4 homo sapien
5	91.5	9.5	384	5 Q9VWJ7	Q9VWJ7 drosophila
6	86.5	9.0	438	5 Q16753	Q16753 caenorhabdi
7	83	8.6	398	16 Q85646	Q85646 escherichia
8	82.5	8.6	1107	4 Q9P2H7	Q9P2H7 homo sapien
9	81	8.4	424	16 Q8Z1G9	Q8Z1G9 salmoneilla
10	80.5	8.4	708	12 Q88442	Q88442 strawberry
11	80	8.3	390	16 Q8XB85	Q8XB85 escherichia
12	80	8.3	390	16 Q8FA92	Q8FA92 escherichia
13	80	8.3	459	10 Q949V7	Q949V7 arabidopsis
14	80	8.3	1139	10 Q8S666	Q8S666 oryza sativ
15	80	8.3	2025	10 Q8LM71	Q8LM71 oryza sativ
16	79.5	8.3	923	16 Q833J4	Q833J4 enterococcu

Q93mw8 mus musculus  
Q94sv1 mus musculus  
Q95gq6 salmoneilla  
Q8pbk2 xanthomonas  
O59773 schizosacch  
Q9lsv7 arabidopsis  
Q7v442 enterobacte  
Q9p3u4 schizosacch  
Q9ym03 broad bean  
Q9fjv3 arabidopsis  
Q90726 human papil  
O52121 escherichia  
Q8liw4 oryza sativ  
Q7sz55 brachydanio  
Q9u013 giardia lam  
Q96h79 homo sapien  
Q8vcz9 mus musculus  
Q9qx62 mus musculus  
Q8va41 chum salmon  
O45377 caenorhabdi  
Q9au13 oryza sativ  
Q8lkw0 oryza sativ  
Q23215 caenorhabdi  
Q9srz7 arabidopsis  
Q8zff9 crimean-con  
Q8t3g9 caenorhabdi  
Q01808 caenorhabdi  
Q9vnt8 drosophila  
Q8d8t2 vibrio vuln

#### ALIGNMENTS

RESULT 1

Q9H7Y0 PRELIMINARY; PRT; 182 AA.  
ID Q9H7Y0  
AC Q9H7Y0;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE Hypothetical protein FLJ14103.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Mammary gland;  
RA Isoqai T., Ota T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Ninomiya K., Iwayanagi T.;  
RT \*NEO human cDNA sequencing project.\*;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK024165; BAB14843.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 182 AA; 20643 MW; CA22BB5607329427 CRC64;  
Query Match 100.0%; Score 962; DB 4; Length 182;  
Best Local Similarity 100.0%; Pred. No. 7.6e-92;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEPQLGPEAAALRPGMLALLVWSALSCFSFLPASSLSLVPOVRTSYNFGRTFLGLDKC 60  
DB 1 MEPQLGPEAAALRPGMLALLVWSALSCFSFLPASSLSLVPOVRTSYNFGRTFLGLDKC 60  
QY 61 NACIGTSICKFFKEEIRSDNWLASHLGLPDSLLSYLPANTYSDSKINRPFVIFRLVSKY 120  
DB 61 NACIGTSICKFFKEEIRSDNWLASHLGLPDSLLSYLPANTYSDSKINRPFVIFRLVSKY 120

```

QY 121 QNEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQGRLEKFLCM 180
Db 121 QNEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQDCHQGRLEKFLCM 180
QY 181 LR 182
Db 181 LR 182

RESULT 2
Q8WX00 PRELIMINARY; PRT; 166 AA.
AC Q8WX00;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE BA435K1.1 (Novel protein) (Fragment).
CN BA435K1.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lawlor S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL591491; CAD13520.1; -.
FT NON TER 166
SQ SEQUENCE 166 AA; 18711 MW; E23F4A20F02E74C1 CRC64;

Query Match 90.2%; Score 868; DB 4; Length 166;
Best Local Similarity 99.4%; Pred. No. 4e-82;
Matches 165; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPOLGPEAAALRPGWLALLWVSALSCSFSLPQVTSYNFGRFTFLGLDKC 60
Db 1 MEPOLGPEAAALRPGWLALLWVSALSCSFSLPQVTSYNFGRFTFLGLDKC 60
QY 61 NACIGTSICKKFFKEIRSDNWLASHLGLPDLSSLYPANYSDSKIMRPVEIFRLVSKY 120
Db 61 NACIGTSICKKFFKEIRSDNWLASHLGLPDLSSLYPANYSDSKIMRPVEIFRLVSKY 120
QY 121 QNEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQ 166
Db 121 QNEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQ 166

RESULT 3
Q8C319 PRELIMINARY; PRT; 435 AA.
AC Q8C319;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE Hypothetical P-loop containing nucleotide triphosphate hydrolases
DE structure containing protein.
GN 4930578C19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT *Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.;
RL Nature 420:563-573 (2002).;
DR EMBL; AK085770; BAC39535.1; -.
DR MGD; MGI:1923155; 4930578C19RIK.

```

```

KW Hypothetical protein.
SQ SEQUENCE 435 AA; 49042 MW; 0A1B46BB04CBB1D CRC64;

Query Match 68.1%; Score 655; DB 11; Length 435;
Best Local Similarity 78.6%; Pred. No. 1.8e-59;
Matches 132; Conservative 8; Mismatches 26; Indels 2; Gaps 2;

QY 1 MEPQ-LGPEAAALRPGWLA-LLWVSALSCSFSLPASSLSLVPOVRTSYNFGRTFLGLD 58
Db 1 MESQWRGAATAFHQHLARLLWVSTLSCSFSLPASSLSLVPOVRTSYNFGRTFLGLD 60
QY 59 KNCACIGTSICKKFFKEIRSDNWLASHLGLPDLSSLYPANYSDSKIMRPVEIFRLVSKY 118
Db 61 KNCACIGTSICKKFFKEIRSDNWLASHLGLPDLSSLYPANYSDSKIMRPVEIFRLVSKY 120
QY 119 KYQNEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQ 166
Db 121 RYQIEISDRRICASVAPKTCISIERILQKTGRFQKWLQAKRLTPDLVQ 168

RESULT 4
Q8NDZ4 PRELIMINARY; PRT; 430 AA.
AC Q8NDZ4;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DE Hypothetical protein.
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC037293; AAH37293.1; -.
KW Hypothetical protein.
SQ SEQUENCE 430 AA; 49481 MW; EB72CACF14E71457 CRC64;

Query Match 12.5%; Score 120; DB 4; Length 430;
Best Local Similarity 26.6%; Pred. No. 0.00068;
Matches 45; Conservative 29; Mismatches 77; Indels 18; Gaps 4;

QY 3 POLGPEAAALRPGWLALLWVSALSCSFSLPASSLSLVPOVRTSYNFGRTFLGLDKCNA 62
Db 7 PKLGRLSRSLKLAALGSLMLVLMVHSPSLASQWRNELT-----DRFLQLNKCPA 57
QY 63 CIGTSICKKFFKEIRSDNW-----LASHLGLPDLSSLYPANYSDSKIMRPVEIFRLV 117
Db 58 CFGTSWCRRELNGQVVFAMGRRLDLDLNVKN-----VYFAQYGRGGRVRVVKRLIG 113
QY 118 SKYQNEISDRKICASAPKTCISIERVLRKTERFQKWLQAKRLTPDLVQ 166
Db 114 SQRELAQLDQSIQCKRATGRCPCDLLQAMPRTFEARLNGDVRLLTPEAVE 162

RESULT 5
Q9VMJ7 PRELIMINARY; PRT; 384 AA.
AC Q9VMJ7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE CG7890 protein (NEO1736p).
GN CG7890.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

```

```

SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA MEDLINE=20196006; PubMed=107311132;
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.A., Bouch J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.I., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina R.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.B.,
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Testor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise B.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan X., Xu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003511; AAF48941.1; -
DR EMBL; AY119100; AAM50960.1; -
DR HSSP; P52848; INST.
DR FlyBase; FBgn0031005; CG7890.
DR GO; GO:0008146; P:sulfotransferase activity; IEA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer; 1.
SQ SEQUENCE 384 AA; 43374 MW; D48346358EA72C59 CRC64;

Query Match 9.5%; Score 91.5; DB 5; Length 384;
Best Local Similarity 24.5%; Pred. No. 0.54; Indels 35; Gaps 5;
Matches 39; Conservative 24; Mismatches 61;

QY 9 AAALRPGWALLLVWYALSLSFSLPASSLSLVPOVRTSYNFGRTFLGLDKCNACIGTISI 68
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 6 ATSAQPAIMVMSICWSSLSRSRLAIKVICVLVYSYF-----NACLLASI 55
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

QY 69 CKXFFKEIRSDNMLASHGLPPDLSLYPANYSDSKWIRVFLRVLYSKYQNEISDRK 128
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 56 SRSLSQSNLR--NLIA-----LTSRPNRTNNA-----ANSSSSSSSRN 94
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

QY 129 ICASASAPKTCSEIRVLKTERF-----QKWLQAKELTPD 163
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

```

Db 370 PQGVKCIIDEMVSALQPRYAASETYLQNT 398

RESULT 8  
Q9P2H7 PRELIMINARY; PRT; 1107 AA.

Q9P2H7  
ID Q9P2H7 PRELIMINARY; PRT; 1107 AA.  
AC Q9P2H7;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein KIAA1370 (Fragment).  
GN KIAA1370.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
R1 NCBI\_TaxID=9606;  
R2 NCBI\_TaxID=9606;  
R3 NCBI\_TaxID=9606;  
R4 NCBI\_TaxID=9606;  
R5 NCBI\_TaxID=9606;  
R6 NCBI\_TaxID=9606;  
R7 NCBI\_TaxID=9606;  
R8 NCBI\_TaxID=9606;  
R9 NCBI\_TaxID=9606;  
R10 NCBI\_TaxID=9606;  
R11 NCBI\_TaxID=9606;  
R12 NCBI\_TaxID=9606;  
R13 NCBI\_TaxID=9606;  
R14 NCBI\_TaxID=9606;  
R15 NCBI\_TaxID=9606;  
R16 NCBI\_TaxID=9606;  
R17 NCBI\_TaxID=9606;  
R18 NCBI\_TaxID=9606;  
R19 NCBI\_TaxID=9606;  
R20 NCBI\_TaxID=9606;  
R21 NCBI\_TaxID=9606;  
R22 NCBI\_TaxID=9606;  
R23 NCBI\_TaxID=9606;  
R24 NCBI\_TaxID=9606;  
R25 NCBI\_TaxID=9606;  
R26 NCBI\_TaxID=9606;  
R27 NCBI\_TaxID=9606;  
R28 NCBI\_TaxID=9606;  
R29 NCBI\_TaxID=9606;  
R30 NCBI\_TaxID=9606;  
R31 NCBI\_TaxID=9606;  
R32 NCBI\_TaxID=9606;  
R33 NCBI\_TaxID=9606;  
R34 NCBI\_TaxID=9606;  
R35 NCBI\_TaxID=9606;  
R36 NCBI\_TaxID=9606;  
R37 NCBI\_TaxID=9606;  
R38 NCBI\_TaxID=9606;  
R39 NCBI\_TaxID=9606;  
R40 NCBI\_TaxID=9606;  
R41 NCBI\_TaxID=9606;  
R42 NCBI\_TaxID=9606;  
R43 NCBI\_TaxID=9606;  
R44 NCBI\_TaxID=9606;  
R45 NCBI\_TaxID=9606;  
R46 NCBI\_TaxID=9606;  
R47 NCBI\_TaxID=9606;  
R48 NCBI\_TaxID=9606;  
R49 NCBI\_TaxID=9606;  
R50 NCBI\_TaxID=9606;  
R51 NCBI\_TaxID=9606;  
R52 NCBI\_TaxID=9606;  
R53 NCBI\_TaxID=9606;  
R54 NCBI\_TaxID=9606;  
R55 NCBI\_TaxID=9606;  
R56 NCBI\_TaxID=9606;  
R57 NCBI\_TaxID=9606;  
R58 NCBI\_TaxID=9606;  
R59 NCBI\_TaxID=9606;  
R60 NCBI\_TaxID=9606;  
R61 NCBI\_TaxID=9606;  
R62 NCBI\_TaxID=9606;  
R63 NCBI\_TaxID=9606;  
R64 NCBI\_TaxID=9606;  
R65 NCBI\_TaxID=9606;  
R66 NCBI\_TaxID=9606;  
R67 NCBI\_TaxID=9606;  
R68 NCBI\_TaxID=9606;  
R69 NCBI\_TaxID=9606;  
R70 NCBI\_TaxID=9606;  
R71 NCBI\_TaxID=9606;  
R72 NCBI\_TaxID=9606;  
R73 NCBI\_TaxID=9606;  
R74 NCBI\_TaxID=9606;  
R75 NCBI\_TaxID=9606;  
R76 NCBI\_TaxID=9606;  
R77 NCBI\_TaxID=9606;  
R78 NCBI\_TaxID=9606;  
R79 NCBI\_TaxID=9606;  
R80 NCBI\_TaxID=9606;  
R81 NCBI\_TaxID=9606;  
R82 NCBI\_TaxID=9606;  
R83 NCBI\_TaxID=9606;  
R84 NCBI\_TaxID=9606;  
R85 NCBI\_TaxID=9606;  
R86 NCBI\_TaxID=9606;  
R87 NCBI\_TaxID=9606;  
R88 NCBI\_TaxID=9606;  
R89 NCBI\_TaxID=9606;  
R90 NCBI\_TaxID=9606;  
R91 NCBI\_TaxID=9606;  
R92 NCBI\_TaxID=9606;  
R93 NCBI\_TaxID=9606;  
R94 NCBI\_TaxID=9606;  
R95 NCBI\_TaxID=9606;  
R96 NCBI\_TaxID=9606;  
R97 NCBI\_TaxID=9606;  
R98 NCBI\_TaxID=9606;  
R99 NCBI\_TaxID=9606;  
R100 NCBI\_TaxID=9606;  
R101 NCBI\_TaxID=9606;  
R102 NCBI\_TaxID=9606;  
R103 NCBI\_TaxID=9606;  
R104 NCBI\_TaxID=9606;  
R105 NCBI\_TaxID=9606;  
R106 NCBI\_TaxID=9606;  
R107 NCBI\_TaxID=9606;  
R108 NCBI\_TaxID=9606;  
R109 NCBI\_TaxID=9606;  
R110 NCBI\_TaxID=9606;  
R111 NCBI\_TaxID=9606;  
R112 NCBI\_TaxID=9606;  
R113 NCBI\_TaxID=9606;  
R114 NCBI\_TaxID=9606;  
R115 NCBI\_TaxID=9606;  
R116 NCBI\_TaxID=9606;  
R117 NCBI\_TaxID=9606;  
R118 NCBI\_TaxID=9606;  
R119 NCBI\_TaxID=9606;  
R120 NCBI\_TaxID=9606;  
R121 NCBI\_TaxID=9606;  
R122 NCBI\_TaxID=9606;  
R123 NCBI\_TaxID=9606;  
R124 NCBI\_TaxID=9606;  
R125 NCBI\_TaxID=9606;  
R126 NCBI\_TaxID=9606;  
R127 NCBI\_TaxID=9606;  
R128 NCBI\_TaxID=9606;  
R129 NCBI\_TaxID=9606;  
R130 NCBI\_TaxID=9606;  
R131 NCBI\_TaxID=9606;  
R132 NCBI\_TaxID=9606;  
R133 NCBI\_TaxID=9606;  
R134 NCBI\_TaxID=9606;  
R135 NCBI\_TaxID=9606;  
R136 NCBI\_TaxID=9606;  
R137 NCBI\_TaxID=9606;  
R138 NCBI\_TaxID=9606;  
R139 NCBI\_TaxID=9606;  
R140 NCBI\_TaxID=9606;  
R141 NCBI\_TaxID=9606;  
R142 NCBI\_TaxID=9606;  
R143 NCBI\_TaxID=9606;  
R144 NCBI\_TaxID=9606;  
R145 NCBI\_TaxID=9606;  
R146 NCBI\_TaxID=9606;  
R147 NCBI\_TaxID=9606;  
R148 NCBI\_TaxID=9606;  
R149 NCBI\_TaxID=9606;  
R150 NCBI\_TaxID=9606;  
R151 NCBI\_TaxID=9606;  
R152 NCBI\_TaxID=9606;  
R153 NCBI\_TaxID=9606;  
R154 NCBI\_TaxID=9606;  
R155 NCBI\_TaxID=9606;  
R156 NCBI\_TaxID=9606;  
R157 NCBI\_TaxID=9606;  
R158 NCBI\_TaxID=9606;  
R159 NCBI\_TaxID=9606;  
R160 NCBI\_TaxID=9606;  
R161 NCBI\_TaxID=9606;  
R162 NCBI\_TaxID=9606;  
R163 NCBI\_TaxID=9606;  
R164 NCBI\_TaxID=9606;  
R165 NCBI\_TaxID=9606;  
R166 NCBI\_TaxID=9606;  
R167 NCBI\_TaxID=9606;  
R168 NCBI\_TaxID=9606;  
R169 NCBI\_TaxID=9606;  
R170 NCBI\_TaxID=9606;  
R171 NCBI\_TaxID=9606;  
R172 NCBI\_TaxID=9606;  
R173 NCBI\_TaxID=9606;  
R174 NCBI\_TaxID=9606;  
R175 NCBI\_TaxID=9606;  
R176 NCBI\_TaxID=9606;  
R177 NCBI\_TaxID=9606;  
R178 NCBI\_TaxID=9606;  
R179 NCBI\_TaxID=9606;  
R180 NCBI\_TaxID=9606;  
R181 NCBI\_TaxID=9606;  
R182 NCBI\_TaxID=9606;  
R183 NCBI\_TaxID=9606;  
R184 NCBI\_TaxID=9606;  
R185 NCBI\_TaxID=9606;  
R186 NCBI\_TaxID=9606;  
R187 NCBI\_TaxID=9606;  
R188 NCBI\_TaxID=9606;  
R189 NCBI\_TaxID=9606;  
R190 NCBI\_TaxID=9606;  
R191 NCBI\_TaxID=9606;  
R192 NCBI\_TaxID=9606;  
R193 NCBI\_TaxID=9606;  
R194 NCBI\_TaxID=9606;  
R195 NCBI\_TaxID=9606;  
R196 NCBI\_TaxID=9606;  
R197 NCBI\_TaxID=9606;  
R198 NCBI\_TaxID=9606;  
R199 NCBI\_TaxID=9606;  
R200 NCBI\_TaxID=9606;  
R201 NCBI\_TaxID=9606;  
R202 NCBI\_TaxID=9606;  
R203 NCBI\_TaxID=9606;  
R204 NCBI\_TaxID=9606;  
R205 NCBI\_TaxID=9606;  
R206 NCBI\_TaxID=9606;  
R207 NCBI\_TaxID=9606;  
R208 NCBI\_TaxID=9606;  
R209 NCBI\_TaxID=9606;  
R210 NCBI\_TaxID=9606;  
R211 NCBI\_TaxID=9606;  
R212 NCBI\_TaxID=9606;  
R213 NCBI\_TaxID=9606;  
R214 NCBI\_TaxID=9606;  
R215 NCBI\_TaxID=9606;  
R216 NCBI\_TaxID=9606;  
R217 NCBI\_TaxID=9606;  
R218 NCBI\_TaxID=9606;  
R219 NCBI\_TaxID=9606;  
R220 NCBI\_TaxID=9606;  
R221 NCBI\_TaxID=9606;  
R222 NCBI\_TaxID=9606;  
R223 NCBI\_TaxID=9606;  
R224 NCBI\_TaxID=9606;  
R225 NCBI\_TaxID=9606;  
R226 NCBI\_TaxID=9606;  
R227 NCBI\_TaxID=9606;  
R228 NCBI\_TaxID=9606;  
R229 NCBI\_TaxID=9606;  
R230 NCBI\_TaxID=9606;  
R

```
RT enterica serovar Typhi CT18." ;
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Ty2 / ATCC 700931;
RC MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18." ;
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL622283; CAD06719.1; -
DR EMBL; ABC16848; AAO1749.1; -
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR001126; UMCUC_like.
DR Pfam; PF00817; IMS; 1.
DR PROSITE; PS01073; UMCUC; 1.
KW Complete proteome.
SQ SEQUENCE 424 AA; 47826 MW; 942565A1381361F5 CRC64;

Query Match      8.4%; Score 81; DB 16; Length 424;
Best Local Similarity 24.5%; Pred. No. 7.5;
Matches 38; Conservative 23; Mismatches 78; Indels 16; Gaps 6;

Qy 23 VSALSCSFSLPAS-----SLSLVPOVTSYNGRTFLGLDKNCAGTSGICKPFKE 75
Db 67 VVAFSSNYELVASMNRVWAHLEELAPRVE-QYSIDEMFLDVRNIDSCIDFDFGRQLRE 125
Qy 76 BIRSDNMLASHLGLPPDLSLYPANYSDSKIWPRVE-IFRLVSKYQNEISDRKICASAS 134
Db 126 VHSSEGLTIGVGMGPTKTLAKSAQWA--SKWRQFEGVLALTS--GNPKRTEKLLSLQP 181
Qy 135 APTCSIERVLRK---TERFQKWLQAKLTDPDLVQ 166
Db 182 VDEIHWGVGRISKKLHTWGIITLQALANPTFFIR 216

RESULT 10
Q88442 ID Q88442 PRELIMINARY; PRT; 708 AA.
AC Q88442;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF v protein.
CS Strawberry vein banding virus.
OC Viruses; Retroid viruses; Caulimoviridae; Caulimovirus.
OX NCBI_TaxID=47903;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98318753; PubMed=9654684;
RA Petrzik K., Benes V., Mraz I., Honetalegrova-Franova J., Ansoerge W.,
RA Spak J.;
RT "Strawberry vein banding virus--definitive member of the genus
RT Caulimovirus." ;
RL Virus Genes 16:303-305(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Petrzik K.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBSJ databases.
DR EMBL; X97304; CA65970.1; -
DR MEROPS; A03.001; -
DR GO; GO:0004190; F:apartic-type endopeptidase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006278; P:RNA dependent DNA replication; IEA.
DR InterPro; IPR000588; Peptidase_A3A.
DR Pfam; PF02160; Peptidase_A3; 1.
DR Pfam; PF00078; rvt; 1.
DR PRINTS; PR00731; CAULIMOPTASE.
```

```
KW RNA-directed DNA polymerase; Transferase
SQ SEQUENCE 708 AA; 81111 MW; A73A8CAC48A47033 CRC64;

Query Match      8.4%; Score 80.5; DB 12; Length 708;
Best Local Similarity 21.7%; Pred. No. 16;
Matches 45; Conservative 31; Mismatches 68; Indels 63; Gaps 9;

Qy 4 QLGPEAAAIRPQWLALLVWSALSC-----SFSU---PA-----SSLSLVQ 43
Db 381 RLAPETIQIQL-----TAFSPQGHYEWLVPFGLKQAPALFORHMDLSLWMYPO 429
Qy 44 VRTSY-----NPGRTFLGLDKNCAGTSGICKPFKEI----- 77
Db 430 PCAVYVDVLIIVFSKTEEBEHLGHVKVILNRCKA-LGIVLSKK--KAQLCKTTFINFLGVIE 486
Qy 78 RSDNWLASHLGLPPDLSLYPANYSDSKIWPRVEIFRLVSKYQNEISDRKICASASAPK 137
Db 487 RGNLKVQSHIGL---HLVAFDPQLSDRNALQRFGLNLYISAVFPKIANLRSPLOVKLKK 543
Qy 138 TCSIERVLRKTERFQKWLQAKLTDPDL 164
Db 544 EITWSWTEKDTETVRKIKSLVKTLEDL 570

RESULT 11
Q8XB85 ID Q8XB85 PRELIMINARY; PRT; 390 AA.
AC Q8XB85;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Orf, hypothetical protein.
GN YJIM OR Z5937 OR ECS5298.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OX Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7." ;
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12." ;
RL DNA Res. 8:11-22(2001).
DR EMBL; AE005665; AAG59520.1; -
DR FIC; B91291; B91291.
DR PIR; D86132; D86132.
KW Complete proteome.
SQ SEQUENCE 390 AA; 43599 MW; B54561646F69415B CRC64;

Query Match      8.3%; Score 80; DB 16; Length 390;
Best Local Similarity 22.1%; Pred. No. 8.6;
Matches 42; Conservative 31; Mismatches 45; Indels 72; Gaps 13;

Qy 39 SLVPOVTSYNGRTFLGLDKC-----NACIGTSGIC---KFFKEIRSDNWLASHLGL 89
Db 39 SLVPOVTSYNGRTFLGLDKC-----NACIGTSGIC---KFFKEIRSDNWLASHLGL 89
```

```
Db 83 NLCFLIKSSVGFQKT-----DKCPYFVSLLVVGTTCDGKKWYE-----YMAE---F 128
QY 90 PPSLLSYPNYSDSK--IWRPVEIFRL-----VSKYQNEIS-----DRKIC 130
Db 129 KPVMVQLNSVKDDASRALWK-AEMRLQKVAEERFGHEISDALDAIALKNRRRAL 187
QY 131 AS-----ASAPKCTSIERVLK-----KTERE-----QKWLQAKRLTPD- 163
Db 188 ANFYHGLQNPPLSGDILKVVYGATFRFDKKEALINELDMATARVQWEEGQRLDPRP 247
QY 164 --LVQDCHOG 171
Db 248 RILITGCPIG 257

RESULT 12
Q8FA92 PRELIMINARY; PRT; 390 AA.
AC Q8FA92;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein yj1M.
GN yj1M OR C5418.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=06:H1 / CF7073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RA "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AF016772; AN83838.1; -.
SQ HYPOTHETICAL PROTEIN; COMPLETE PROTEOME.
KW SEQUENCE 390 AA; 43611 MW; C12BD2BFED7E83C8 CRC64;

Query Match 8.3%; Score 80; DB 16; Length 390;
Best Local Similarity 22.1%; Pred. No. 8.6;
Matches 42; Conservative 31; Mismatches 45; Indels 72; Gaps 13;

QY 39 SLVPQVTSYNGFTFLGDKC-----NACIGTSIC---KFPKKEIRSDNMLASHLGL 89
Db 83 NLCFLIKSSVGFQKT-----DKCPYFVSLLVVGTTCDGKKWYE-----YMAE---F 128
QY 90 PPSLLSYPNYSDSK--IWRPVEIFRL-----VSKYQNEIS-----DRKIC 130
Db 129 KPVMVQLNSVKDDASRALWK-AEMRLQKVAEERFGHEISDALDAIALKNRRRAL 187
QY 131 AS-----ASAPKCTSIERVLK-----KTERE-----QKWLQAKRLTPD- 163
Db 188 ANFYHGLQNPPLSGDILKVVYGATFRFDKKEALINELDMATARVQWEEGQRLDPRP 247
QY 164 --LVQDCHOG 171
Db 248 RILITGCPIG 257

RESULT 13
Q949V7 PRELIMINARY; PRT; 459 AA.
AC Q949V7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN AT3G59300.

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RA "Full length cDNA of gene F25L23 160/AT3G59300 (GI:7801680).";
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.; Clones.";
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
DR EMBL; AY050863; AAK92800.1; -.
DR EMBL; AY091205; AAM14144.1; -.
KW Hypothetical protein.
SQ SEQUENCE 459 AA; 51592 MW; B7E100293E5BF436 CRC64;

Query Match 8.3%; Score 80; DB 10; Length 459;
Best Local Similarity 31.3%; Pred. No. 10;
Matches 41; Conservative 13; Mismatches 53; Indels 24; Gaps 8;

QY 23 VSALSCSESLPAS-----SLSLVPQVTSYNGFTFLGDKNACIGTSIC-KFP 73
Db 2 VIAAASPSLGSCHQSYDEFSIPYKRTNARNRVDFGCGSAMLSSRCKLPFF 61
QY 74 KEIRSDNMLASH-LGLPPLSLLSYPNYSD---DSKI--WRPVEIFRLSKYQNEISDR 127
Db 62 GSAPFVSS--GGHDLGL---TKVSVADYSDSVDFSGFYGHPLDL----KPSKRVQET 112
QY 128 KICASASAPKT 138
Db 113 KLSASEVARTT 123

RESULT 14
Q8S666 PRELIMINARY; PRT; 1139 AA.
AC Q8S666;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSJNEA0034B05.27.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Sasaki C., Henry D., Oates R., Simmons J.;
RA "Rice Genomic Sequence.";
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
```

```
Db 83 NLCFLIKSSVGFQKT-----DKCPYFVSLLVVGTTCDGKKWYE-----YMAE---F 128
QY 90 PPSLLSYPNYSDSK--IWRPVEIFRL-----VSKYQNEIS-----DRKIC 130
Db 129 KPVMVQLNSVKDDASRALWK-AEMRLQKVAEERFGHEISDALDAIALKNRRRAL 187
QY 131 AS-----ASAPKCTSIERVLK-----KTERE-----QKWLQAKRLTPD- 163
Db 188 ANFYHGLQNPPLSGDILKVVYGATFRFDKKEALINELDMATARVQWEEGQRLDPRP 247
QY 164 --LVQDCHOG 171
Db 248 RILITGCPIG 257

RESULT 12
Q8FA92 PRELIMINARY; PRT; 390 AA.
AC Q8FA92;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein yj1M.
GN yj1M OR C5418.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=06:H1 / CF7073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RA "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AF016772; AN83838.1; -.
SQ HYPOTHETICAL PROTEIN; COMPLETE PROTEOME.
KW SEQUENCE 390 AA; 43611 MW; C12BD2BFED7E83C8 CRC64;

Query Match 8.3%; Score 80; DB 16; Length 390;
Best Local Similarity 22.1%; Pred. No. 8.6;
Matches 42; Conservative 31; Mismatches 45; Indels 72; Gaps 13;

QY 39 SLVPQVTSYNGFTFLGDKC-----NACIGTSIC---KFPKKEIRSDNMLASHLGL 89
Db 83 NLCFLIKSSVGFQKT-----DKCPYFVSLLVVGTTCDGKKWYE-----YMAE---F 128
QY 90 PPSLLSYPNYSDSK--IWRPVEIFRL-----VSKYQNEIS-----DRKIC 130
Db 129 KPVMVQLNSVKDDASRALWK-AEMRLQKVAEERFGHEISDALDAIALKNRRRAL 187
QY 131 AS-----ASAPKCTSIERVLK-----KTERE-----QKWLQAKRLTPD- 163
Db 188 ANFYHGLQNPPLSGDILKVVYGATFRFDKKEALINELDMATARVQWEEGQRLDPRP 247
QY 164 --LVQDCHOG 171
Db 248 RILITGCPIG 257

RESULT 13
Q949V7 PRELIMINARY; PRT; 459 AA.
AC Q949V7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN AT3G59300.

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RA "Full length cDNA of gene F25L23 160/AT3G59300 (GI:7801680).";
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.; Clones.";
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
DR EMBL; AY050863; AAK92800.1; -.
DR EMBL; AY091205; AAM14144.1; -.
KW Hypothetical protein.
SQ SEQUENCE 459 AA; 51592 MW; B7E100293E5BF436 CRC64;

Query Match 8.3%; Score 80; DB 10; Length 459;
Best Local Similarity 31.3%; Pred. No. 10;
Matches 41; Conservative 13; Mismatches 53; Indels 24; Gaps 8;

QY 23 VSALSCSESLPAS-----SLSLVPQVTSYNGFTFLGDKNACIGTSIC-KFP 73
Db 2 VIAAASPSLGSCHQSYDEFSIPYKRTNARNRVDFGCGSAMLSSRCKLPFF 61
QY 74 KEIRSDNMLASH-LGLPPLSLLSYPNYSD---DSKI--WRPVEIFRLSKYQNEISDR 127
Db 62 GSAPFVSS--GGHDLGL---TKVSVADYSDSVDFSGFYGHPLDL----KPSKRVQET 112
QY 128 KICASASAPKT 138
Db 113 KLSASEVARTT 123

RESULT 14
Q8S666 PRELIMINARY; PRT; 1139 AA.
AC Q8S666;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSJNEA0034B05.27.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Sasaki C., Henry D., Oates R., Simmons J.;
RA "Rice Genomic Sequence.";
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
```



Search completed: June 8, 2004, 16:25:48  
Job time : 43 secs